

Perceived risk in online services and its effect on password strength

A quantitative study investigating risk perception in online services
to determine its effects on password strength

HÅKON ORVIK BRENDE
OLE-MARTIN MÅNSSON

SUPERVISOR

Jaziar Radianti
Lucia Castro Herrera

University of Agder, 2022

Faculty of Social Science
Department of Information System

Master

Preface

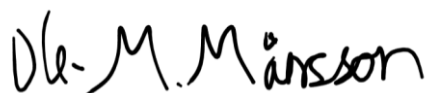
This master's thesis marks the end of the master's degree in Cybersecurity at the University of Agder 2022. It has been two rewarding years with lots of new knowledge, understanding, and challenges.

The topic of password security is interesting and relevant due to its continuous presence in daily life, and the lack of understanding of users' behavior. The topic has been rewarding to work with, and we hope our findings provide relevant insight into why users behave as they do, and which services are more exposed than others.

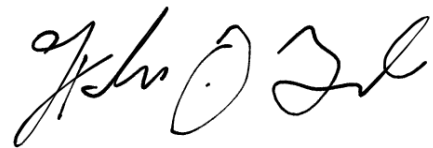
The completion of this thesis was possible due to the continued support, help and guidance from our supervisors. We want to thank Associate Professor Jaziar Radianti, and PhD Research Fellow Lucia Castro Herrera. Thank you for your time dedication and availability when we needed your guidance, and your constructive feedback and positive attitudes throughout the whole thesis. We would also like to thank all the participants partaking in our study. This master's thesis was possible thanks to all of you.

Kristiansand

03.06.2022



Ole-Martin Månsson



Håkon Orvik Brende

Abstract

Passwords are the most used method for authentication in online platforms. At the same time, password management continues to be one of the biggest security risks for individual users. This is due to both inadequate password behavior of most users, especially related to password strength which depends on the parameters assigned by the user in most cases. Two of the most prevalent behaviors that can expose users to danger are password reuse and weak password strength.

Our thesis focuses on the problem of weak password strength usage. Therefore, we seek to answer the following research problem: “*Does the perception of risk associated with different online services influence password strength and is this universally applied?*”. We conducted the research study on Norwegian students from the University of Agder. To answer our question, we followed a quantitative methodology in form of an online distributed survey. The study was based on findings from a literature review which helped us get an understanding of different factors affecting users’ password behavior, risk perception, knowledge, and the state of password strength. The survey received 99 respondents of which 70 were eligible for further analysis. The analyses of the data were conducted using Excel. We present our findings in figures, tables, and descriptive analysis. Our results show that using different password strengths for different online services is common among users. In addition, there are no significant changes in password strength between the services when analyzing behaviors of individual users. Moreover, the perceived risk of user accounts being attempted compromised, and the consequences of compromise in services have low correlation with password strength, with a few exceptions for some services. Two of these exceptions being porn, and news. Furthermore, we discuss our findings in detail by looking at outliers and trends in the data, and some commonalities between the services that follow a similar pattern in our findings. We concluded that password strength differs among online services, and that certain online services are more likely to have weaker passwords than others.

Table of contents

	Preface.....	3
	Abstract	4
1	INTRODUCTION	9
	1.1 Motivation	10
	1.2 Research gap.....	11
	1.3 Previous Work.....	12
	1.4 Problem statement	13
	1.5 Thesis structure.....	14
2	LITERATURE REVIEW	16
	2.1 Introduction	16
	2.2 Searching for literature.....	16
	2.3 Reviewing the literature	20
	2.4 Findings	21
	2.4.1 Password strength	21
	2.4.2 Knowledge	23
	2.4.3 Perceived risk.....	24
	2.5 Summary	27
3	METHODOLOGY.....	29
	3.1 Research approach.....	29
	3.2 Research design.....	30
	3.3 Population and sample	31
	3.4 Survey design	31
	3.4.1 Limitations	32
	3.4.2 Variables	32
	3.5 Analysis	36
	3.5.1 Correlation	36
4	RESULTS	37
	4.1 Initial results	37
	4.2 Demographic	38
	4.3 Risk perception.....	41
	4.3.1 Password strength	42

4.3.2	Likelihood of compromise	43
4.3.3	Importance of availability	44
4.3.4	Consequences	45
4.4	Comparing men and women	46
4.5	Summarize results	49
5	DATA ANALYSIS	51
5.1	Users	51
5.2	Risk factors	53
5.3	Online services.....	55
5.3.1	Educational	55
5.3.2	Email.....	55
5.3.3	Financial	56
5.3.4	Gaming	56
5.3.5	News	56
5.3.6	Porn.....	57
5.3.7	Single purpose	57
5.3.8	Shopping.....	58
5.3.9	Social media	58
5.3.10	Video streaming.....	58
5.3.11	Transportation.....	59
5.3.12	Travel.....	59
5.3.13	Work	59
6	DISCUSSION.....	61
6.1	RQ1. Password strength.....	61
6.2	RQ2. Perceived risk in digital services	62
6.3	RQ3. Behavior Patterns	63
6.4	Limitations	64
7	CONCLUSION	66
7.1	Final remarks	66
7.2	Future work.....	66
8	REFERENCES	67
	APPENDIX	71

Appendix A Survey	71
-------------------------	----

List of figures

Figure 1	Research design	30
Figure 2	Distribution of education amongst the participants	38
Figure 3	Password strength in online services	42
Figure 4	Perceived likelihood of compromise	43
Figure 5	Importance of availability	44
Figure 6	Perceived consequences.....	45
Figure 7	Comparing password strength distribution between men and woman.....	46
Figure 8	Comparing perceived likelihood of compromised between men and women	47
Figure 9	Comparing importance of availability between men and woman	48
Figure 10	Comparing the perceived consequences of a compromise between men and woman.....	49
Figure 11	Relationship between mean password strength and standard deviation.....	52
Figure 12	Correlation between password strength and likelihood of compromise.....	53
Figure 13	Correlation between password strength and importance of availability.....	54
Figure 14	Correlation between password strength and consequences	54

List of tables

Table 1	Literature Review Inclusion Criteria	19
Table 2	Literature Review Exclusion Criteria	19
Table 3	Inclusion and Exclusion Progression.....	21
Table 4	Concept Matrix	27

Table 5	Overview of online services	34
Table 6	Demographic Information	38
Table 7	Average scores per education with highest values outlined in bold. Results that are not statistically significant are greyed out.	39
Table 8	Responses to negative experiences and password managers	40
Table 9	Data breaches effect on password strength and perceived risks	40
Table 10	Password managers effect on password strength and perceived risks	41
Table 11	Correlating age and risk variables	51

1 INTRODUCTION

Passwords are the most used user authentication method as it has remained the go-to form of end-point authentication for several decades and is to this day still considered the de-facto king due to its frequent and easy usage (Robinson, u.d.; Vekua, 2021; Seitz, Pfab, & Souque, 2017). It is to this day mostly reliant on end-users' behavior as to how much protection this method provides and is prone to human error and other human shortcomings. It therefore continues to be a main concern for security experts for the unforeseeable future due to the rising threats in cybersecurity, and the unlikelihood of an alternative authentication systems replacement passwords as the most used method of user authentication (Taneski, Heričko, & Brumen, 2016).

New studies, annual reports and national cyberthreat-reports continue to prove that we are still utilizing less than desirable password habits such as reusing passwords across multiple accounts sharing passwords with someone else, and using simple, predictable passwords such as birthdates, keyboard strokes and dictionary words, making passwords a poor form of protection against the ever-increasing number of threats (Verizon, 2022; Nilsen, 2020; LastPass, 2019; PST, 2022).

Despite knowing that passwords are there for the reason of authorizing end-users, protect their information's confidentiality, people are still inclined to neglect passwords hygiene across multiple accounts. Many users will also use the same password for both critical services such as email, and non-critical services such as social media (Security.org Team, 2021; Google, 2019; SpyCloud, 2021). There seems to be two general categories of explanations. The first explanation is that they don't understand the risks involved. It is shown that people do not fully understand who an attacker might be, or the methods an attacker could utilize (Salem, Moreb, & Rabayah, 2021). Users also tend to overestimate how much protection their security measures and passwords provide (Ur, et al., 2016). The second category of explanation is that users do understand the risks and consequences but choose bad password habits due to convenience (Gratian, Bandi, Cukier, & Dykstra, 2018). This is extremely prevalent in the case for password reuse, as users struggle to remember unique passwords. Password sharing being another example of convenience over consequence (Merdenyan B. P., 2019).

These habits of bad password management such as reuse, weak password, etc, is still under investigation as researchers and cyber professional are still trying to

understand the reasoning behind the behaviors to better mitigate its security breach potential (Awad, 2016; Merdenyan B. P., 2019; Merdenyan & Petrie, 2022). This is done by explaining the limitations of the human mind to singlehandedly create, store, and apply an ever-increasing number of passwords. There have been attempts to find whether there are underlying factors influencing our decision making on how we perceive risk, benefit, or gain in a specific way. The cultural differences aspect could also influence our behavior as language both written and oral can be factors in password strength and habits (Grobler, et al., 2020). How password habits are linked to personality traits (Gratian, Bandi, Cukier, & Dykstra, 2018). Using fear to better users' decisions (Rodríguez-Priego, Bavel, Vila, & Briggs, 2020). How knowledge influence our decisions (Salem, Moreb, & Rabayah, 2021). These are just some aspects of the current literature out there, indicating that the password "problem" does not have a simple explanation. Since it is difficult to holistically understand why we're doing what we're doing when it comes to passwords. Here is where we found a gap in the literature, which fails to look at how different online services affect end-users' behaviors.

1.1 Motivation

Our primary motivation for researching this topic is our interest in how users perceive and use passwords. We have realized that we ourselves as soon-to-be cybersecurity-experts have fallen to some of the "pitfalls" when managing our passwords, contributing to the mediocre practices we would warn against. After some inquiries about password habits in a small pre-study we did prior to starting this thesis. In reading reports, and evaluating our own behavior, we have confirmed our suspicion that this security risk is more widespread than first assumed. We wanted to investigate the topic and find a gap in the literature where we could add new knowledge and more understanding of the problem. We noticed that passwords had been researched for many years by different cybersecurity scholars, and that it could be hard to find a gap in the literature. Even so, we started generally searching the topic two months before starting this study. At that time, we were researching different factors affecting password behaviors, and even performed a small literature review on the subject. Our findings during that literature review were that the human's problem with memory, fear factors, ease of use, password creation strategies, risk, attitude, and knowledge, were some of the factors that had been studied. We also found that risk had several different ways of affecting password habits, and thought it had potential for further investigation in our master thesis. This is when we started the literature review process for our master thesis.

We were specifically researching risk perception and the effects it has on users' passwords. This is when we found a gap in the literature and decided to make it our study and thesis. In other words, our personal interest and the gap in the literature were the two factors that motivated us to finally pursue the topic and make it our master thesis.

The main goal of this study is to get a better understanding of how users' password strength is distributed amongst different online services. The second goal is to investigate whether there exists a correlation between password strength and perception of risk associated with online services, and that password strength is not uniformly distributed amongst the services. We do this by investigating how end users', such as higher education students, self-reported password strength on different online services. Then we correlate password strength with how users perceive the importance of availability, the likelihood of compromise, and the consequences of compromise for each service. By doing this we want to find out if there are noticeable trends or patterns in our collective password usage. We discover whether end users are prone to lacking in security in some online services in comparison to others. We are also given a better understanding of how perceived risk is correlated with these trends.

1.2 Research gap

There have been done extensive research on the topic of passwords as authentication since its first application. A search test using only "password" as a keyword returned on Scopus and Google Scholar there were 16.457 and over 5.000.000 articles respectively as of 1st of February 2022, proving that passwords have been under extensively researched and thoroughly scrutinized for the past decades, and probably will be for the next decades to come.

Researchers have come to terms that password are most likely here to stay as they are easy to implement, "provides little friction to users workflow", and are already the most common and well-known method of user authentication (451 Research; ADSelfServicePlus). Researchers are therefore trying to find out how to mitigate the damage potential of bad password habits by studying and understanding why we are doing what we are doing and to what degree. Several studies have investigated which risks users know of and if they are aware of what are considered as bad and good password-related behaviors (Ur, et al., 2016; Salem, Moreb, & Rabayah, 2021). These are often specific risks such as password sharing or what constitutes a secure password. Previous research has not addressed broadly how users perceive the likelihood of someone trying to compromise their

account, and how severe the consequences of a compromise would be. This is more akin to how a security employee would determine the level of risk, and how we are exploring users' perception of risk. We do this to get a better understanding of how these two factors affect users' password strength.

Past studies have also not investigated how accounts for different online services would affect the users' perception of risk in this way. The only similar study we found was Merdenyan and Petrie. They investigated end-users' perception of risks of password-related activities, concluding that end users are viewing password behaviors differently in association with risks, thinking for instance that anything password related with e-Banking is considered riskier in comparison to social networking, email and eCommerce, proving that we're not viewing risk uniformly across digital services (Merdenyan & Petrie, 2017). One study investigated risk perception as a driver for netizens (internet users) cybersecurity behavior, indicating that affect heuristics could be a major influencer for our perception, as affect, risk and benefit, and how this could impact the cybersecurity domain (Schaik, Renaud, Wilson, Jansen, & Onibokun, 2020).

In this thesis, we therefore study the correlations between perceived risk, account-type importance, and users' self-reported password strength for several distinct categories of services by relying on descriptive statistics after our review of the existing academic literature. There is a need to explore how end-users perceive different online services in relation to likelihood of compromise and consequence of compromise. Our goal is to further explore how these perceived risks are correlated to password strength. To our knowledge, there has been little research and few scholars who has investigated this topic in this way. We thereby acknowledged it as a research gap.

1.3 Previous Work

This study is based on our previous work we have conducted in other courses a few months prior to the master thesis. It consisted of a literature review which gave the impression that there were gaps in the research related to the distribution of passwords strength across multiple services. There was also a small mixed study while consisted of a small quantitative survey and short interviews. The survey was conducted to verify if there were any significant differences in perception of different online services. The survey (n=15) gave a skewed perception as some services clearly outperformed others regarding risk, benefit, and consequences. To further investigate the results, we conducted a semi-constructed interview with two

students confirming our suspicion that password strength is not uniformly distributed amongst online services.

1.4 Problem statement

Users use weak passwords, we do not know which online services are more exposed than others, and past studies have not investigated how dividing risk into perceived likelihood of compromise and consequence of compromise affects password strength.

One way security risks can occur is when a great number of people are committing the same or similar patterns of neglect in password management. The purpose of this thesis is therefore to investigate whether there are any indications or coherent patterns in end-user's password habits regarding their distribution of passwords strength across multiple digital services. This will hopefully reveal weaknesses in users' collective password usage and make the already existing risks visible. Certain sectors of digital services could be more vulnerable for compromise in comparison to other digital services. Not being able to distinguish such trends is most likely due to the nature of passwords inherent secrecy, limiting end-users' knowledge of similar use-cases; allowing this potential pattern to remain unknown. We're attempting to answer this by firstly disproving the notion of uniform password strength distribution being a common practice, find differences in perceived risks associated with digital services and analyze the correlation between these factors. The goal is to disclose any potential security threats by shedding light on any pattern indicating possible vulnerabilities, as well as encouraging further research about password habits. To address this potential security risk, we're going to research the following question: *Does perception of risk associated with different online services influence password strength, and is this universally applied?*

More specifically, we will answer the following research questions:

RQ1: To what degree is uniform password strength distributed amongst end-users?

RQ2: What is the correlation between password strength and perceived risk in digital services?

RQ3: To what degree are end users consistent in their behaviors when choosing password strength for different online services?

The first research question is important as it used to show whether or not end-users apply similar passwords strength for different online services. The purpose is to show that end-users do not treat online services equally, and that some sectors

of online services are more exposed than others. The second research question is to specifically focus on how perceived risk affects password strength. This is one of the identified gaps in the literature we are trying to answer as password strength and perceived risk of online services has to our knowledge not been explored in the literature. The third research question is to explore how a user differ in their answers on password strength for each service. The purpose is to show on the aggregate how much online services affects the behavior of each individual end-user.

1.5 Thesis structure

This section consists of an overview of the thesis organization, i.e., the order of how our study was conducted from the theoretical background, all the way to the conclusions of our study.

Chapter 2. Literature review. This chapter contains our literature review methodology. It explains our inclusion and exclusion criteria. The search process for findings literature in details, and the way we performed the review process. The findings from the literature review are presented in a concept centric method, with a summary at the end to clearly state what has been done before, and what we are doing differently.

Chapter 3. Research approach. Describes the research approach used to answer the research question. This includes the choices made, how the quantitative methodology was used in the study and presents how the survey was formed and distributed.

Chapter 4. Results. Presenting the overall findings of the survey, doing a preliminary examination of the results and the different perceptions of digital services before doing the correlative analysis. We also discuss if our sample group is representative of the target population we want to study.

Chapter 5. Analysis and findings. Conducting correlation analysis between password strength and perceived risks in different online services, investigating user behaviors to find patterns of habits and other variables to investigate factors influencing password strength. Online services will be summarized as individual categories where key-findings will be presented.

Chapter 6. Discussion. We discuss the research questions and our findings. We also compare our findings with the existing literature to either expand upon, support or contradict previous studies. We also discuss some limitations and potential biases with our research approach.

Chapter 7. We present the conclusions of the study, and some potential ways of further researching the topic and expand upon what we found.

2 LITERATURE REVIEW

2.1 Introduction

The purpose of the literature review is to find previous studies on perceived risk and knowledge effects on users' password strength and other password related habits. This is done to identify a research gap where our study can contribute with more understanding of users' risk perception and password strength habits. We follow the systematic literature review method of Webster and Watson (Webster & Watson, 2002). This method combines rigor while allowing for flexibility during the literature process using concepts to structure the review and findings. We describe the search process which explains how we systematically found the literature, including the inclusion and exclusion criteria. The review process which explains how we decided on which articles to include, and at the end is a summary of our findings and key takeaways and concept matrix.

2.2 Searching for literature

Following the Webster and Watson method, we started looking for articles in highly regarded journals before searching the academic research databases. First, we focused on collecting articles from the most prominent journals in information systems (IS) research, the basket of eight. We manually searched each of the journals one by one, and our only search criteria was 'password' with no restrictions. This was because there were not that many results while searching for more specific terms such as 'password risk' or 'password behavior'. We also did not want to restrict the time frame, as we wanted a holistic picture on password research from its origins. After searching "password" in each basket of eight journal, we screened the title of every result for relevant topics. We also investigated the abstracts of the articles with titles that seemed to be relevant and could fit our problem statement. We ended up with one relevant article after searching thru each of the journals. Which suggests that research on password has not been a core interest in information systems. We are unsure as to why there are so few articles about passwords in the basket of eight articles. One of the possible explanations being that it is hard to investigate passwords due to privacy concerns

and ethical concerns. Which could result in studies struggling to reach the quality to get published in one of these journals. We also thought that relevance could be one of the reasons as the basket of eight journals focuses on new theories and unexplored topics. However, we would then expect to find lots of articles. Maybe even dating two to three decades back. Another explanation could be that they have not received a study which had strong enough evidence supporting their new hypothesis and theory. As we will present in our findings section, it does seem that there are no one major explanation as to why users behave as they do.

Next, we conducted an automated search on the academic database Scopus. We started searching for password habits and password behavior in the titles, abstract or keywords. This gave us numerous results, a notable contrast to our experience with searching the basket of eight journals. As most of the articles were not relevant to our area of concern, we leveraged a more advanced search criteria by including the word “risk” to our search. Searching for password, behavior, and risk in the title, abstract or keywords gave us 151 results. We were hoping to get more results and decided to try one more search string before starting to screen the titles. We therefore modified the search criteria where “risk” only had to be mentioned anywhere. We also included in the search string that the study had to mention either behavior or habit, no longer restricting it to only behavior. The search string looked like this:

(TITLE-ABS-KEY(password) AND TITLE-ABS-KEY(behavior OR habit) AND ALL(risk)).

This search resulted in 362 articles which was manageable to search and sort. We were also aware that behavior could be spelt both as “behavior” and “behaviour”. However, the spelling did not change anything and had no impact as to which results were shown. As we did with the basket of 8 articles, we screened each article first by title and then abstract. The result was 27 potentially relevant articles out of the 362.

After our first round of searching, we realized that we preferred to know which articles were peer-reviewed and published in journals, and which were not. We then had to copy paste the titles of each of the 27 articles to the search string in Scopus and added the “published in journal” criteria to the search string. We searched all of the 27 articles and started to sort out journal articles from non-journal articles. Searches that did not yield a result were double checked by reading and searching within the article itself for a mention of journal publication. We also copy pasted the title of that article on Google Scholar to see if it was published in a peer reviewed journal somewhere. If we were not able to find it in any journal, we classified it as a “non-journal article” and went on to the next article and did the same process. Searches that yielded a result were checked by clicking on the

result and reading the title and some of the abstract, to ensure that this result was the same as the article we were searching for. Thus, out of 27 articles, 17 articles were published in academic journals and 10 articles from conferences papers.

At this point we started a first round of review. During the review process we found two new relevant articles as they were cited in two other articles. This method is referred to as snowballing, it is performed by investigating potentially relevant articles that have been cited and referenced in other articles (Webster & Watson, 2002). We now had 29 articles for further analysis.

We wanted to find more articles relating to risk perception after our initial round of search and review. Therefore, following the same title and abstract screening procedure, we applied two different search strings. The first being “TITLE-ABS-KEY (password) AND TITLE-ABS-KEY (risk AND awareness)” in journals or conferences. It provided us with 59 results, five of them being new potentially relevant articles. We also searched “TITLE-ABS-KEY (password) AND TITLE-ABS-KEY (risk AND perception)” in journals or conferences. That gave us 60 results, of which seven of them being potentially relevant articles. We were not yet sure at that point if we had discovered any of these 12 new articles before in our first round of searching. We would not realize if they were new additions or not before reviewing them.

Our inclusion and exclusion criteria were based on our problem statement and research questions. This made it easier for us to decide which articles were potentially relevant when reading the titles and abstracts during the search process. Which also helped us narrow down the articles to a manageable amount that could be manually read and analyzed. Some of the criteria could not be enforced during the search process as it required us to review the article to determine whether the article followed the criteria or not. This is why we ended up with several articles during the search process which ended up being excluded after reading the methodology, results, and findings of the article during the review process. We provide the criteria and the reasoning for each criterion below.

Table 1 Literature Review Inclusion Criteria

Criteria	Justification
The study must focus on passwords	Our study investigates passwords and password strength. Studies on other forms of authentication such as biometric authentication, pin codes, multi factor authentication, or certificate-based authentication, are not relevant to our study. Passwords are unique in that the users themselves are most often responsible for the strength of the security measure.
The study has to contain information about user behavior	This is a broad criterion, because as long as the study tells us something about users' behavior it will be included. This means that the studies do not have to focus on factors that affect user behavior, but could just be descriptive fact about user behavior. Simply describing how users act in some kind of way is enough to be included. It is an important criterion however as it does exclude articles that only focus on technical aspects. For example, articles that only focus on defining what a strong password is or focus on defining what good password behaviors are.
The study must be written in English or Norwegian	It would not be possible for us to understand the articles otherwise.

Table 2 Literature Review Exclusion Criteria

Criteria	Justification
Studies that focus on single sign on	We are not investigating single sign on in our study. How single sign on affects password strength or other user behavior is therefore not relevant to include in the literature review.
Studies that focus on humans' problem with password memorability	The memorability problem is often mentioned in studies investigating factors affecting users' password behaviors. We are however not investigating how humans' insufficient memory capabilities are affecting password strength. We
Studies that focus on password creation strategies	We do not investigate how different password creation strategies affect password strength and find these studies irrelevant to our study.

2.3 Reviewing the literature

We started reading through the articles in alphabetical order when sorting by titles. We made an excel sheet containing an empty concept matrix. As we read thru the articles, we started plotting in concepts and article titles in the matrix. We marked relating concepts and articles with an X in the matrix. We noted down sample size, data collection method, relevant results, and some data analysis in a word file as we read thru the articles. This was especially important for the articles that were not published in journals. Our reason being that articles published in journals are peer reviewed when searching in the Scopus database (Scopus, u.d.). Journal articles are therefore less likely to overpromise in their conclusions.

We excluded five of the 27 articles after the first round of review process. One article was excluded as we found out when reading that it had nothing to do with passwords and only mentioned it in the title. Four more articles were also excluded due to not being relevant and failing the inclusion criteria. Then, following the "snowballing" method proposed by Webster and Watson (Webster & Watson, 2002), while looking at our articles we also looked at the references for relevant articles. at referenced articles if we found relevant citations while reviewing the articles. This meant that the search process was ongoing. We added two new articles during our initial review following the snowball method as they were referenced in other articles. That gave us 24 articles after the first round of review.

We initially ended up with 12 concepts after our first reading of the 24 articles. After some discussion and refocusing our research problem. We narrowed it down to 3 concepts. Mostly because several of the first concepts were different password habits such as sharing, storing, reusing, etc. Those were not relevant for our study as we are not investigating how these behaviors are affected by different online services or risk perception. This is also where we discovered that we wanted to find more literature on "perceived risk and risk awareness,". That's when we started our second round of search in Scopus, and reviewed the potentially relevant articles just like before, which is detailed above in the "search" chapter. We ended up with two more articles after reading the 12 articles we found in the second round of review. Most of the relevant articles were already in our library of relevant articles. This also gave us more confidence in that we had found most of the relevant articles. That gave us a total of 26 articles in our literature review, after reading thru the 41 articles we found during our search process. The next step was then to determine how to present our findings and further analyze the articles. Table 3 summarizes our inclusion and exclusion progress and process.

Table 3 Inclusion and Exclusion Progression

Source	Inclusion and Exclusion Process			
	Added articles during search	Excluded articles during review	New Inclusions	Total Included Articles
Basket of eight	1	0	1	1
First round: Scopus	26	-5	21	22
Articles	2	0	2	24
Second Round: Scopus	12	-10	2	26

2.4 Findings

This is where we present the three concepts, password strength, knowledge, and perceived risk, based on the 26 articles. Throughout this section, we explain the takeaways and key findings of the articles. We note where there is directly or indirectly conflicting results, and discuss our own understanding of the conflicts. This is also true for findings in one study which could either be directly or indirectly supported by other studies. Each concept is explained, defined, and given an example as it is mentioned in the text. Each paragraph has its own angle as to how the concept is explored. For example, one paragraph is focused on how sex is explored and studied within the *perceived risk* concept. Then the next paragraph focuses on how age is explored in the same concept.

2.4.1 Password strength

We chose this concept to show and establish that it is still relevant to investigate what affects password strength and that users' passwords strength continues to be too weak. 9 out of the 26 articles mentioned password strength. Either as an entropy in form of a quantified number of guess a password could withstand, or general strength judgements on users' passwords. We should mention that all of the studies used different ways of determining entropy as there are several different ways of brute force guessing a password. The studies also used different reference points when performing strength judgments and when determining the strength of a password. We do not see this as a problem however as they all used very similar methods of determining what is a strong or weak password. What would have been considered a strong password in one study would in almost all cases been considered a strong password in the other studies due to their use of similar criteria. The consensus is that people still use weak passwords, and this is supported by several different studies including field experiments, a leaked password dataset analysis, and a meta-analysis on the literature which concluded that humans still

use weak passwords (Taneski, Heričko, & Brumen, 2019; Juozapavicius, Brilingaite, Bukauskas, & Lugo, 2022; Grobler, et al., 2020).

One dataset analysis study from 2022 focused on differences between genders and ages based on a dataset from 2018 - 2021. They found that the overall password strength corresponded to a randomly six-character generated password, even though the average length was nine characters (Juozapavicius, Brilingaite, Bukauskas, & Lugo, 2022). This is also supported by a much bigger dataset analysis from 2021 which gave similar results with shorter length and lower entropy (Grobler, et al., 2020). Most likely due to the dataset including older accounts and consisting of many different smaller breaches. We're also not sure what the password rules were for the different services, which also could affect the results. A longitudinal study from 2017 also support these findings (Renaud & Zimmerman, 2017).

The dataset analysis study from 2022 also found that the difference in password strength between the genders was significant in every age group (Juozapavicius, Brilingaite, Bukauskas, & Lugo, 2022). Male users on the aggregate used passwords almost twice as hard to guess as females (Juozapavicius, Brilingaite, Bukauskas, & Lugo, 2022). This finding could also be indirectly supported by another study from 2018 which studied users' intentions to use strong "password generation". Password generation combines both the intentions to use a strong password, and the intention to not reusing passwords (Gratian, Bandi, Cukier, & Dykstra, 2018). However, there are some inconsistencies in the literature on how intentions affect behavior. The gender differences are also contested by two other studies. One field experiment study from 2016, and one other survey analysis from 2018 (Steinbart, Keith, & Babb, 2016; Cain, Edwards, & Still, 2018).

One study from 2011 also investigated university students and how year of study was correlated with password strength. They found that "first year students created weaker password than third year students." (Tarwireyi, Flowerday, & Bayaga, 2011). They also found IT students had the strongest passwords.

One predictor of strong passwords were the personality trait "conscientiousness" and avoidant decision-making style (Gratian, Bandi, Cukier, & Dykstra, 2018).

One 2005 study looked at users' password strength differences by interviewing IT "experts", and surveying end users (Stanton, Stam, Mastrangelo, & Jolton, 2005). They found also that the experts seemed to have stronger passwords in general, but there were big outliers.

2.4.2 Knowledge

There were 13 studies that investigated how users' knowledge affect password habits. Knowledge is simply defined as what users perceive to know, and how this perceived knowledge relates to the real world. For example, users' think that a password should withstand more than 1000 guesses, but in the real world this number is much higher depending on if it is an online or offline attack. The studies investigates knowledge of potential risks, what constitutes a strong password and good password habits, attack methods, potential attackers, etc. This concept is heavily built on articles from conference papers as there was not many relevant articles about the concept we defined as knowledge in journals.

There have been different findings on age. One 2020 study found no difference in knowledge between age groups (Rodríguez-Priego, Bavel, Vila, & Briggs, 2020).

One study from 2015 focused on the difference between low- and high web literacy users and found that users with low web literacy reported worse password habits (Rinn, Summers, Rhodes, Virothaisakun, & Chisnell, 2015). That included password sharing, password strength, and password reuse (Rinn, Summers, Rhodes, Virothaisakun, & Chisnell, 2015). This finding between low- and high web literacy users supports the 2005 study of IT experts which found that the experts in general used stronger passwords (Stanton, Stam, Mastrangelo, & Jolton, 2005). The meta-analysis from 2019 also found that users lacked knowledge about their password habits (Taneski, Heričko, & Brumen, 2019). These findings were even more prevalent amongst low web literacy users (Rinn, Summers, Rhodes, Virothaisakun, & Chisnell, 2015). One study from 2010 however found that most users understood what constituted a strong password but did not get any results indicating that knowledge affected their behavior (Tam, Glassman, & Vandenwauver, 2010). Rather that willingness to sacrifice convenience was the predictor of actual behavior. Three other studies also found that knowledge was not enough to change actual behavior, and that most users have enough knowledge to act better than they do (Cain, Edwards, & Still, 2018; Tarwireyi, Flowerday, & Bayaga, 2011; Fredericks, Futcher, & Thomson, 2016).

When looking at how gender correlates with knowledge, it seems that males have more knowledge about cyber hygiene and password management (Cain, Edwards, & Still, 2018; Salem, Moreb, & Rabayah, 2021). However, as with a 2010 and 2011 study, they did not find any difference in actual behavior between the genders in 2018 either (Cain, Edwards, & Still, 2018; Tarwireyi, Flowerday, & Bayaga, 2011; Tam, Glassman, & Vandenwauver, 2010). Based on these, knowledge does not seem to change behavior. One study from 2020 however

found that people with more knowledge self-reported better password habits, and as with the other studies found that men had “higher levels of knowledge of secure passwords” (Kennison & Chan-Tin, 2020, s. 8). Another study also found that security experts had stronger passwords than non-experts (Creese, Hodges, Whitty, & Jamison-Powell, 2013). Men used stronger passwords than females (Juozapavicius, Brilingaite, Bukauskas, & Lugo, 2022). Both studies suggesting that knowledge does affect certain behaviors.

When comparing knowledge of different password habits, the literature shows that users have little knowledge about the danger of reusing passwords, especially when compared to the knowledge of creating strong password (Salem, Moreb, & Rabayah, 2021). However, when it comes to specifically explain what makes a password stronger than other passwords there are miss conceptions (Ur, et al., 2016). Users also does not know how many guesses a password should withstand. 34% participants said a password should withstand less than 50 guesses. 67% said less than 50,000 guess (Ur, et al., 2016).

A study from Polen found that students got their knowledge from the internet or friends instead of what the article called “qualified sources of knowledge” (Szumski, 2018, s. 1277).

2.4.3 Perceived risk

We found 15 articles that investigated users’ perception of different risks. These could be risks associated with different types of user behaviors, specific accounts, user information, attackers, etc. Perceived risk will be based on a users’ knowledge (the previous concept), but risk in this context is more in form of a subjective measurement or scale. Knowledge is simply restating what users know or perceive to know. One example of knowledge is: A user knows it’s bad to reuse passwords. One example of perceived risk is: The user thinks reusing passwords is a little bad, but using weak passwords is much worse.

A 2022 study claims that users does not consider the risk from entire datasets being leaked (Juozapavicius, Brilingaite, Bukauskas, & Lugo, 2022). This could be part of the explanation as why users do not know why reusing passwords are such a bad idea. Users do not perceive the risks of not changing passwords, only the benefits (Merdenyan & Petrie, 2022; Merdenyan B. P., 2019). The 2022 study also found this to be the case with password reuse as well. However, another study found that the users did perceive the risks with password reuse but that the benefits outweighed them (Merdenyan & Petrie, 2017). This is also the case with shared accounts and passwords. Users seem to perceive the risks, but the benefits

outweigh the risks (Merdenyan & Petrie, 2022; Merdenyan B. P., 2019). The 2022 study also found users do perceive the risks with storing passwords.

One study from 2016 found that “password management” is considered the riskiest behavior (Parsons, Butavicius, McCormac, & Calic, 2016). They did not compare it to other security behaviors, only other general internet usage behaviors such as social networking, which also was perceived to be very risky.

Another study showed that specifying warning messages was more effective than general warnings. For example, focusing on presenting a financial loss risk warning, and then what to do to reduce risk and avoid loss (Rodríguez-Priego, Bavel, Vila, & Briggs, 2020). This is supported by another study which found that risks perceived to have immediate negative consequences led to better behavior (Tam, Glassman, & Vandenwauver, 2010). The same study also say that “Users must feel a personal loss if the account is compromised” (Tam, Glassman, & Vandenwauver, 2010, s. 242). That certain risks motivate better than others. Lastly, that privacy risks were the main motivator for good password habits (Tam, Glassman, & Vandenwauver, 2010). We can conclude that when users have a specific negative outcome in mind it will affect behavior. In low literate computer users however, perceived consequences did not seem to affect behavior (Rinn, Summers, Rhodes, Virothaisakun, & Chisnell, 2015). The researchers speculate that the participants though their current behavior to be secure enough due to lack of knowledge (Rinn, Summers, Rhodes, Virothaisakun, & Chisnell, 2015).

Trust in a vendor increased “bad decisions” (Rodríguez-Priego, Bavel, Vila, & Briggs, 2020). Trust in others is also linked to password sharing in adolescents (Ouytse, 2021).

There were no differences in self-reported risky behaviors between the genders (Kennison & Chan-Tin, 2020).

Teenagers seems to perceive great risks but underestimate their vulnerability. The same study says “Perceived severity, and fear of cyber threat do not influence teenagers’ compliance intentions” (Mwagwabi & Jiow, 2021, s. 12).

Exposing users to news stories about security breaches “rapidly motivated protective behavioral responses”, stronger passwords being one of the changes (Mamonov & Benbunan-Fich, 2018, s. 40).

Affect valence has an impact on risk perception, and there is a “correlation between risk- and benefit perception (Schaik, Renaud, Wilson, Jansen, & Onibokun, 2020). The benefits of sharing a computer with the same password vs the risks of sharing a computer with the same password.

Both experts and non-experts provide “similar risk assessments” when rating different behaviors (Creese, Hodges, Whitty, & Jamison-Powell, 2013). The same study also found a negative correlation between the experts and non-experts when

looking at perceived risk and password strength. Not sure how well this supports the conclusion that experts use stronger passwords than non-experts due to perceived risk. It could also just be the knowledge and attitude towards security in experts that lead to this behavior.

Users do not think the risk differences between different domains (SNS, email, eBanking, eComm) are big, even though eBanking was considered the riskiest (Merdenyan & Petrie, 2017). Their study also found that sharing passwords with a trusted one was considered the least risky “bad behavior”. They also found a negative correlation between perceived risk and real life behavior but did not find any correlation between perceived consequences and real life behavior.

One study focused on personality traits, risk taking, and internet behavior. They found that financial risk-taking and health/safety risk-taking were “unique predictors of strong password generation” (Gratian, Bandi, Cukier, & Dykstra, 2018).

2.5 Summary

Table 4 Concept Matrix

Contribution	Concepts		
	Password Strength	Knowledge	Perceived Risk
Juozapavicius, et al. (2022)	x		x
Stanton, et al. (2005)	x	x	
Pattinson, et al. (2016)			x
Steinbart, et al. (2016)	x		
Rodriguez-Priego, et al. (2020)		x	x
Cain, et al. (2018)	x	x	
Ouytsel. (2021)			x
Grobler, et al. (2021)	x		
Taneski, et al. (2019)	x	x	
Rinn, et al. (2015)		x	x
Tam, et al. (2010)		x	x
Schaik, et al. (2020)			x
Merdenyan & Petrie. (2022)		x	x
Kennison & Chan-Tin. (2020)			x
Mwagwabi & Jiow. (2019)			x
Mamonov & Benbunan-Fich. (2018)			x
Gratian, et al. (2018)	x		x
Renaud & Zimmerman. (2017)	x		
Salem, et al. (2021)		x	
Creese, et al. (2013)		x	x
Tarwireyi, et al. (2011)	x	x	
Ur, et al. (2016)		x	
Szumski. (2018)		x	
Fredericks, et al. (2016)		x	
Merdenyan & Petrie. (2017)			x
Merdenyan & Petrie. (2019)			x

The concept matrix above makes it easy to see that password strength and perceived risk has little overlap in prior research. The table also shows that most of the articles we were able to find were published during or after 2016. Could be because we were searching in online academic libraries. Older articles may not have been published online. Passwords have been around since the 1960s (McMillan, 2012). We would therefore expect to find articles dating decades back, even though we did not.

To sum up our findings: Previous literature that investigated password strength often focused on users' password entropy and guess ability, confirming that users do not use strong passwords. This shows as that password strength is still a valid concern worthy of exploration. We also found that most research that investigated

user knowledge and its correlation with password strength focused on what users considered to be a secure password, attack methods, and potential attackers. The research showed different findings as to how these correlated with password strength, and it seems that knowledge is only a small factor that affects password strength. Several articles had conflicting findings as to whether or not knowledge even affected real life behaviors. We conclude based on our analysis of previous studies on risk and knowledge as effects on user behavior that there is a need to further study risk perception as a function of password strength. We therefore made this our focus of the thesis.

Only one study tries to answer how online services affect password strength and risk perception, and it misses some things which we will explain. As they note in their limitations: "we did not ask people what negative experiences they had had in relation to password activities, but respondents' answers to the open-ended questions showed that this can have a very strong effect on risk perception" (Merdenyan & Petrie, 2017, s. 8). This is one point we will try to answer. We also investigate more subcategories of online services, as their survey only had four big categories. We would argue that their categories of services were too big, and it does not ultimately help online services with understanding which of them is more exposed. They also did not focus on password strength which we will do. Their focus was password storing and sharing (Merdenyan & Petrie, 2017).

Another important aspect which differentiates our approach is that past studies asked participants what they thought was risky or what the risk was for certain activities. As we are security focused, we quickly realized that we wanted to divide the risk definition up into two questions to explore which part of the risk aspect is affecting users. That is the reason why we are not directly asking participants to rank the level of risk associated with the different online services. We found one study which asked participants to rate the consequences of different activities such as sharing passwords, etc. They found no correlation between perceived consequence of that behavior and real-life behavior (Merdenyan & Petrie, 2017). This study concluded that users know that a behavior is risky, but as they have never experienced any negative consequences associated with that behavior, they rank the consequences as low. Their example being sharing a password with a colleague. Most users seem to consider this risky, but do not perceive the negative consequences (Merdenyan & Petrie, 2017). Therefore, we believe that how perceived likelihood of someone trying to compromise an account, and how the perceived consequences of an account compromise are correlated with user behavior.

3 METHODOLOGY

The research questions from chapter 1 can be answered in multiple ways and the following research approach must be evaluated according to several factors. Two of our main concerns when evaluating the research approach was time and money due to the time constraint of six months. The second most important factor was how it would be able to answer the research questions. The different approaches have several positive and negative aspects to them, and we discuss these in the next paragraphs.

3.1 Research approach

Qualitative approach is used to gather in-depth insight into a given topic by accumulating non-numerical data through video, audio, or text. This approach is often used to “... *collect open-ended, emerging data with the primary intent of developing themes from the data*” (Creswell, 2003, p. 18). The research questions could possibly be answered by conducting one-on-one interviews with relevant participants in the population, organizing a focus group, or distributing a survey with open ended questions. This would enable us to understand end users’ behavior through an open-ended survey (Bhandari P. , Qualitative research, 2022), and motivation and reasoning when perceiving online services and its password in an interview. This approach was discarded because it would not make generalize for a wider population, which we seek to do in our research questions

Quantitative approach is used to analyses numerical data to find trends and patterns between different variables to make generalization in a population. The data are usually collected through experiments, surveys, or observations (Bhandari P. a., 2021). It is possible with a large enough sample size to find statistical indications or correlations which could quantify and to a certain degree find indications for a wider population.

It was decided that our study was best fitted to a quantitative approach. The reason being that our investigation is mainly focused on verifying the existence of the phenomenon instead of understanding why it potentially exists. As we already knew which variables, we wanted to include because of our earlier literature review findings, it was not necessary to further receive elaborations from interview

subjects, lessen the need to pursue new inputs on variables. This is also one of the reasons why we chose to not include open ended questions with possibilities for self-written answers. We chose to conduct a survey as it “*it provides a numeric description of trends, attitudes, or opinions of a population by studying a sample from the population*” (Creswell, 2003, p. 153) which was exactly what we were after.

3.2 Research design

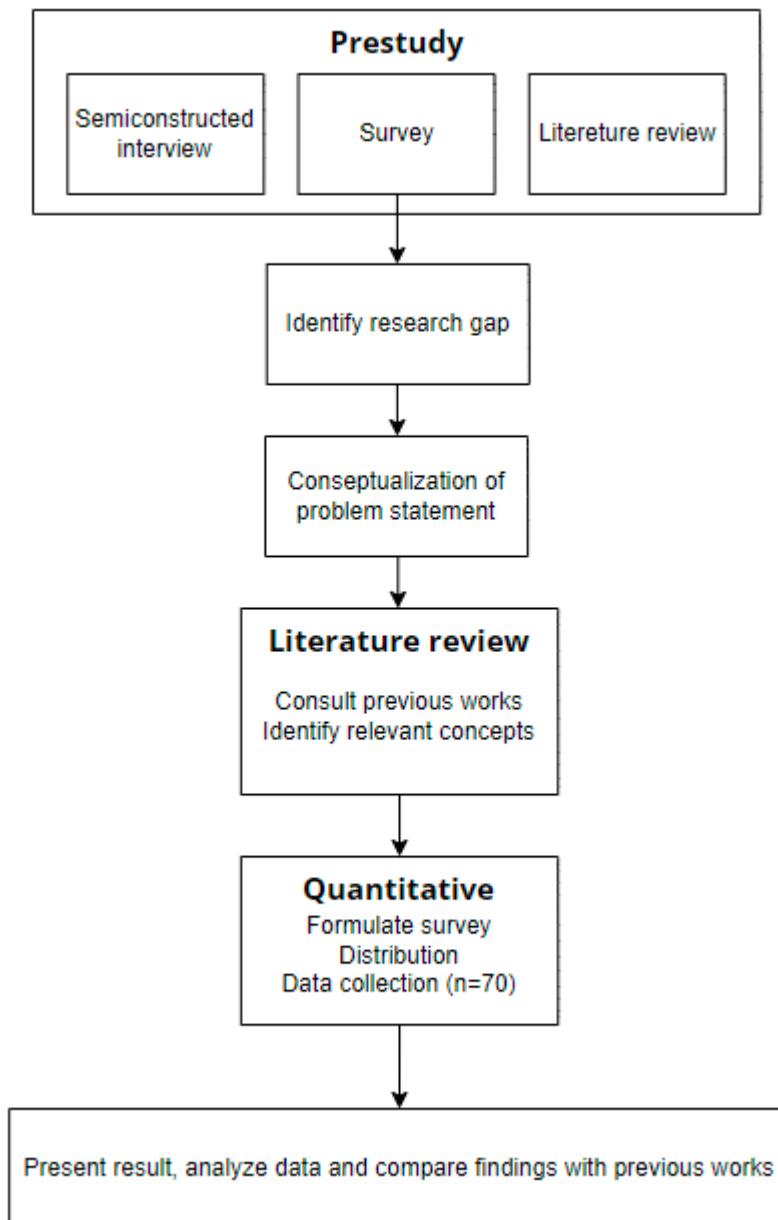


Figure 1 Research design

We will use a descriptive and correlational design. The concepts we defined in the literature review will be measured against a constructed list of general online services through a quantitative survey. The survey will collect data from the targeted population and produce a dataset of descriptions of attitudes toward different online services. This dataset will be presented in a straightforward fashion as is (Lambert, 2012) and analyzed to find correlations between password strength and risk factors, and by doing this we will answer our research questions.

3.3 Population and sample

The research questions could be answered by different populations, but due to the limitation in resources and time it was decided to focus on students. Ideally this would include multiple universities across Norway to achieve a broader distribution. It was more realistic to only include students from the University of Agder (UiA) because the survey would only be available for 2-3 weeks, and it would most likely be unevenly distributed in UiA favor because of our presence at UiA.

The population of UiA exceeds 13,000 students (UiA, 2022). It is not feasible for us to achieve such a high number with the resources at our disposal, so we've calculated a more realistic estimate. If we limit ourselves to a confidence level ranging from 90-95%, a 10% margin of error, and applying it to the formula first for the unlimited population $n = \frac{z^2 * p(1-p)}{\epsilon^2}$ and then by the finite population $n' = \frac{n}{1 + \frac{z^2 * p(1-p)}{\epsilon^2 N}}$, where z is the z-score, ϵ is the margin of error, N is the population, and p is the population proportion (Israel, 1992; Calculator, 2022). We're given the following sample range of 68 to 95 respondents to aim for.

3.4 Survey design

The survey is a method used to collect data from a sample representing a larger population. It involves a series of questions which are usually answered by multiple-choice options which translate to numerical values (Bhandari P. , Scribbr, 2022). The survey is created in SurveyXact as it was recommended by the university, and as opposed to Google Forms it does not log IP addresses and makes it easier to keep the survey anonymous – as we wanted to comply with the GDPR requirements. Thus, we also did not include questions related to personal information.

The online survey was then exported to an XML-format (Microsoft Excel-format). This made it easy for us to export, clean, analyze and visualize the data and findings using Microsoft EXCEL.

The next section will further explain the demographic variables, testing parameters that were included in the survey, as well as the selected testing parameter, and selections of online services commonly used by the target group under study, to answer the research questions.

3.4.1 Limitations

The data was collected at one point in time, with a nonprobability sample where the participants were chosen by convenience (Creswell, 2003, p. 156) because it was not possible to conduct a random sample at the time of the surveys' distribution.

Due to limited resources, as well as the time limit of the masters' thesis, we were forced to accept a less than desirable number. The distribution of the survey was mostly directed at students of the same academic discipline due to their level of proximity to the researchers, giving a skewed distribution of participants where some academical disciplines are overrepresented and some are not represented at all.

Investigating a research gap always involves the element of uncertainty. The lack of concrete knowledge on people's perception of their own password strength regarding different digital services presents us with the possibility that there are other major factors which should have been included or a combination of variables which would have yielded significant results.

3.4.2 Variables

We chose to include demographic variables such as sex, age, completed years of higher education and what kind of education they are currently studying. These were included respectively to validate whether our sample of respondents were representative of the population of the university. We also wanted to investigate whether sex, age, and field of education provided different responses. As mentioned in the literature review, past studies have provided conflicting findings as to how sex correlate with password strength and risk perception (Juozapavicius, Brilingaite, Bukauskas, & Lugo, 2022; Steinbart, Keith, & Babb, 2016; Cain,

Edwards, & Still, 2018). We also suspect that the different fields of education would look at risks and password strength differently.

Included in the survey are some questions related to experiences and the use of technology support. Asking if the participants have any prior negative experiences in the form of being hacked or subject to a data compromised, and if they are currently using a password manager. We believe that negative experiences could potentially have positive effects as it might function as a wake-up-call to improve password habits with the intention of prevent similar incidents to occur and comparing those who had such experiences with those that have not regarding password strength. Studies have shown that those who relied on technologies to manage passwords tends to have both stronger and more unique password. By asking if the participants are currently using a password manager, we'll be able to re-confirmed similar studies, and verify our own results.

3.4.2.1 Online services

As there is no precedence to our knowledge for a set of services which is applicable to a large population which is universally acknowledged and used, we were forced to create our own list of online services as similar studies usually only use the most generic categories, such as social media, online banking, and email. To achieve a general understanding of the population's password behavior not in general sense but segmented to different aspects of their password habits to online services from private to public, and important to unimportant.

To further investigate this, a list of digital services was created which encompasses a broad use of digital services that we expect the participants to be able to relate to, many if not all. There are of course several other categories of accounts what could have been included, such as devices like smartphones and personal computers, health related services, insurance etc, but these were excluded because they either don't use alphanumeric (letters, numbers, and special characters) passwords, but rather personal identification number codes, more commonly known as PIN, or they have the same login option on multiple services. This is especially true in Norway as Norwegians use MinID for over 1000 services from the government (Difi, n.d.).

The list consists of 14 different online services, even though some are similar (travel and transport) in nature they are all unique in their purpose, and some may be more widespread and frequently used in comparison.

Table 5 Overview of online services

Digital service	Examples
Finance	Online banking, investments
Education	School platforms, online courses
Email	Outlook, Gmail, Hotmail
Gaming	Steam, Epic Games, Xbox Game Pass
Music	Spotify, Tidal, Apple Music
News	Online newspapers, news outlets
Pornography	Adult entertainment
Shopping	Known as eCommerce, buy things online, receive in mail
Single purpose	Zip, online drawing tools etc
Social media	Facebook, Instagram, Twitter, LinkedIn, Snapchat
Video streaming	Common form of entertainment. Netflix, YouTube
Transportation	Daily transportation: car, bus, train
Travel	Vacation: hotels, flights
Work	Services which require a company login

Most of the online services used in this thesis are commonly used and we perceive them to be applicable to many if not all the participants. We chose to include pornography as a category, as it is to our knowledge underrepresented in the field of cybersecurity as it rarely if at all included in articles and studies related to risk perception and, in our case, password strength. It stands out compared to other online services as it could contain sensitive information in the form of sexual orientation and practices. This is information which could potentially inflict harm in the wrong hands in the form of blackmail, and it could potentially yield interesting results to gauge the participants responses.

There are other online services that could be included, but the list is only comprised of services which requires a user-account with an adjacent password. we are not testing specific digital services like LinkedIn, Facebook, Office 365, YouTube and airline companies, because people are using and viewing them in different ways. O365 could just as likely be considered either work- or school-related depending on the person asked – the same could be said for email. We also chose not to group services together in categories due to expected nuances in participants answers, and example of this grouping could be gaming, music, porn, and video streaming as entertainment, this could potentially give a too generalized answers to a rather diverse genre.

3.4.2.2 Risk variables

Our four testing parameters were passwords strength, importance of availability, likelihood of compromise, and consequence of compromise. We used the 5-point Likert scale for each question about the perception and behavior towards a given parameter. This meant that each question had five different possible answers. The answers were scaled in a linear fashion, and such that the participants had to rate their responses between two opposing standpoints. Two general examples of this being, strongly agree all the way to strongly disagree, or very high to very low. This gave us the ability to gage the respondent's attitude towards the parameters (Joshi, 2015). It also made the results easy to analyze as each possible response could be assigned a numerical value.

Password strength. Asking participants to rate their different passwords regarding strength across multiple services, but since studies have shown people to have difficulties deciding what passwords are strong (Ur, et al., 2016). As it is difficult to get the participants to submit their current, active passwords for an objective measurement, we're limited to a theoretical approach – and thus invite potential biases associated with self-reporting, where participants could potentially report what they would rather do, instead of what they do. Password strength is indicative to how well protected a user-account is and will be used as a baseline of comparison.

Importance of availability. To investigate if a password protects something which is perceived as important has any impact on its strength. The participants were asked to gage how important each corresponding user-account is. To not misinterpret the question, a brief explanation was given to let the participants have a uniform understanding. As importance is a very subjective term, we associated the importance variable with availability.

Likelihood of compromise. Asking the participants to rate how they perceived likelihood of someone attempting to compromise their user-accounts for each service. Assuming people perceive the likelihood to vary depending on the service or the information stored.

Consequences of compromise. It should be the most obvious factor influencing password strength as it a leading cause for compromise (Verizon, 2022), and as logic would dictate that the greater consequences warrant greater password strength. How severe the consequences of a compromise would be for the participant is ranked from insignificant to significant.

3.5 Analysis

The dataset would be exported to XML-format and cleaned prior to any analysis, and following the descriptive and correlative design, we will first present the data in a coherent format, then conduct further statistical analyzes.

3.5.1 *Correlation*

Correlation ranges from -1 indicating a 1:1 negative correlation, to 1 indicating a 1:1 positive correlation. If you get a correlation of 1 between two variables, it means that when variable x either increases or decreases, variable y will always follow, either increasing or decreasing with variable x . If you get a correlation of -1 between two variables, it means that when variable x either increases or decreases, variable y will always do the opposite, either increasing or decreasing in opposition with variable y . Correlation does not tell us how much the variables move together or in opposition, only how often they move together or in opposition.

Parameters would then be tested against other parameters, password strength, perceived likelihood of compromise, how beneficial each service is, and severity of consequences would be correlated against each other to calculate a correlation coefficient which could be interpreted by comparing it to a range of values with associated descriptions. We will also be investigation user behaviors regarding demographic values.

Correlation is also not capable of telling us anything about causation between variables.

4 RESULTS

This chapter will present the results from the survey prior to the analysis, which will be presented in chapter 5. The data will show the participants demographic distribution used to verify reliability as to the sample's representation of the population. An overview will be given of the respondent's answers regarding the previously mentioned online services to password strength, perceived likelihood of compromise, importance of availability, and consequences.

4.1 Initial results

The survey was active for the duration of 2.5 weeks, accumulating 99 respondents. After cleaning the dataset, removing outliers and half-finished respondents which apparently started but stopped too early to be significant. Those who answered questions beyond the demographic questions, contributing to some of the testing parameters were included in the analysis. For those who completed the survey, but due to either errors or intent submitted unlikely answers; like their age being either 3 or 100 years old - those valuables were omitted from the calculation as they would heavily influence the results. The dataset also contained additional errors which occurred during the surveys creation, where SurveyXact gave incremental values for each change in either question or answer for it to have a unique identifier - resulting in some questions valuing linear answers in a non-linear fashion, this was corrected by manually allocating correct numerical values to the affected variables.

The result after cleaning the dataset was 70 participants eligible for analysis. This is just above the minimum requirement of 68, as defined in chapter 3.3, and still sufficient for our purposes (Israel, 1992). Cronbach Alpha was used to measure the dataset's reliability by calculating the datasets internal consistency, and its value was calculated to 0.989 which was sufficient for further analysis.

4.2 Demographic

For demographic data of respondents, we collected information on gender, aged and completed years of higher education.

Table 6 Demographic Information

Demographic Information	Values
Gender	Male: 52 (72.0%) Female: 17 (24.3%) Others: 1 (1.4%)
Age	20 - 39 years Mean: 26.6 years
Completed years of higher education	1-10 years Mean: 4.4 years

Most of the participants turned out to be male (72.0%), which was expected as computer science and engineering; the most represented fields of study amongst the participants is predominantly male. Whereas the female (24.3%) and those who preferred not to answer (1.4%). The age of the participants is within the expected range of university students, with a few outliers. Ranging from 20 to 39 years old with the majority in their mid-late twenties. The average age is 26.6 years old.

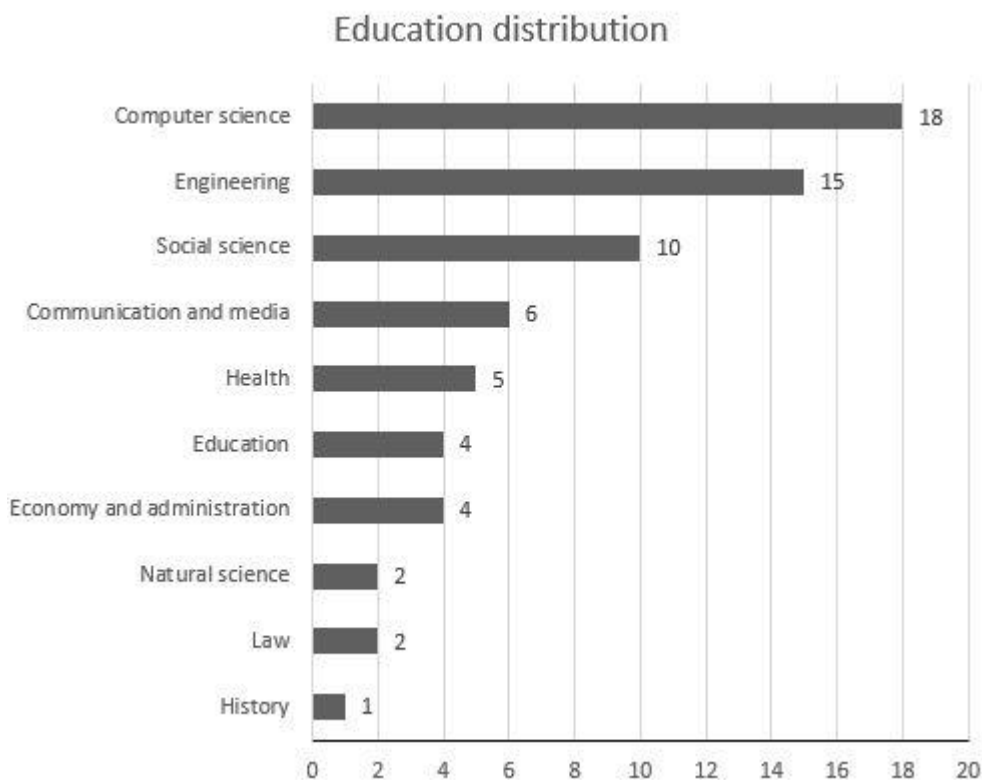


Figure 2 Distribution of education amongst the participants

Table 7 Average scores per education with highest values outlined in bold. Results that are not statistically significant are greyed out.

n	Education	Passwo rd	Likeliho od	Importan ce	Consequen ce
1	History	3.56	3.25	4.45	3.69
2	Natural science	3.73	3.00	3.31	3.38
2	Law	3.17	2.73	4.58	4.00
4	Education	3.68	3.46	3.99	3.52
4	Economy and administration	3.15	3.16	3.39	3.84
5	Health	3.15	3.14	3.96	3.13
6	Communication and media	3.48	2.98	3.97	3.42
10	Social science	3.69	3.48	4.08	3.54
15	Engineering	3.64	3.33	3.87	3.66
18	Computer science	3.91	2.79	3.59	3.23

In this table we have broken down the dataset to present how the different educational fields reports their password strength, perceived likelihood of compromise, importance of availability, and consequences. The data is present in mean value, and values in bold indicate the highest value. Those who did not specify their education were omitted in this table.

When comparing the educations with n=10 or more respondents, there are only small differences in overall password strength, as there is only 0.27 value separating the highest and lowest indicating that by this small sample size it is difficult to dictate whether education has any influence over password strength. Investigating the same for perceived likelihood of compromise paints a slightly broader picture as there is a wider difference (0.69) between the highest and lowest average between the different educations. Similar findings are also present in importance of availability as there it has 0.49 difference between the highest and the lowest. Variation is present at consequences, but only a smaller difference of 0.43 between highest and lowest. Because of the small sample size, we could assume that the differences are negligible enough to influence further analysis in chapter 5. A similar study has been done to measure guessability of passwords for an entire university, and by using the education of Fine Arts as a baseline were able to see differences between educations and password guessability, there results

showed indication of computer science students having the strongest passwords, followed by science and engineering (Mazurek, et al., 2013).

The sample consists of an uneven distribution of students from different educations. As computer science and engineering is overrepresented in this sample, there are also other faculties and educations which is not represented at all. Taking the education into account it makes sense that sex distribution is also skewed in favor of men as Technology and science consists of 66% male and 34% female (UiA, 2022).

Table 8 Responses to negative experiences and password managers

Have you been hacked or subjected to a data breach	Yes: 26 (39%) No: 27 (40%) Don't know: 14 (21%)
Are you using a password manager	Yes: 29 (39%) No: 40 (54%) Previously: 2 (3%)

There were similar results from those who have experienced being hacked or breached as those who had not. The severity of the hack/breach is not quantified in this survey, which means it could range between a few hits at haveibeenpwned.com to being subjected to ransomware and extorted for money. The majority did not use a password manager to store their credentials, only 39% said they did.

Table 9 Data breaches effect on password strength and perceived risks

	Password strength	Likelihood of compromise	Importance of availability	Consequences
Yes	3.835	3.228	3.733	3.378
No	3.642	2.976	3.920	3.437
Don't know	3.603	3.336	3.794	3.786
Δ high/low	0.232	0.360	0.187	0.407

This table shows the average of password strength, and the average of the different risk values when the participant was asked if they have been subjected to a data breach or have previously been hacked. This shows a little effect on either the password strength or risk factors. Most noteworthy was that those who have experienced data breach reported slightly stronger passwords. Those who do not know if they have been subjected to a data breach believes the likelihood of compromise and consequences is highest. Ignorance is apparently not bliss.

Table 10 Password managers effect on password strength and perceived risks

	Password strength	Likelihood of compromise	Importance of availability	Consequences
Yes	3.903	3.071	3.854	3.407
No	3.555	3.190	3.838	3.578
Previously	3.708	3.346	3.250	3.308
Δ high/low	0.347	0.276	0.604	0.270

This table shows the average of password strength and the different risk values when the participant was asked if they are using a password manager. Those who do, and those who did have slightly stronger password compared to those who don't, which was expected as studies have found that those who relies on technical support like password managers have stronger passwords (Lyastani, Schilling, Fahl, Bugiel, & Backes, 2018).

4.3 Risk perception

The graphs presented here shows only the respondents answers which gave a satisfactory answer, as those who reported a service as not applicable (N/A) was not included the following graphs. Results are sorted for each category by answers, presenting the digital services in descending order starting from the ones with the highest percentage of the highest scoring variable to the lowest.

4.3.1 Password strength

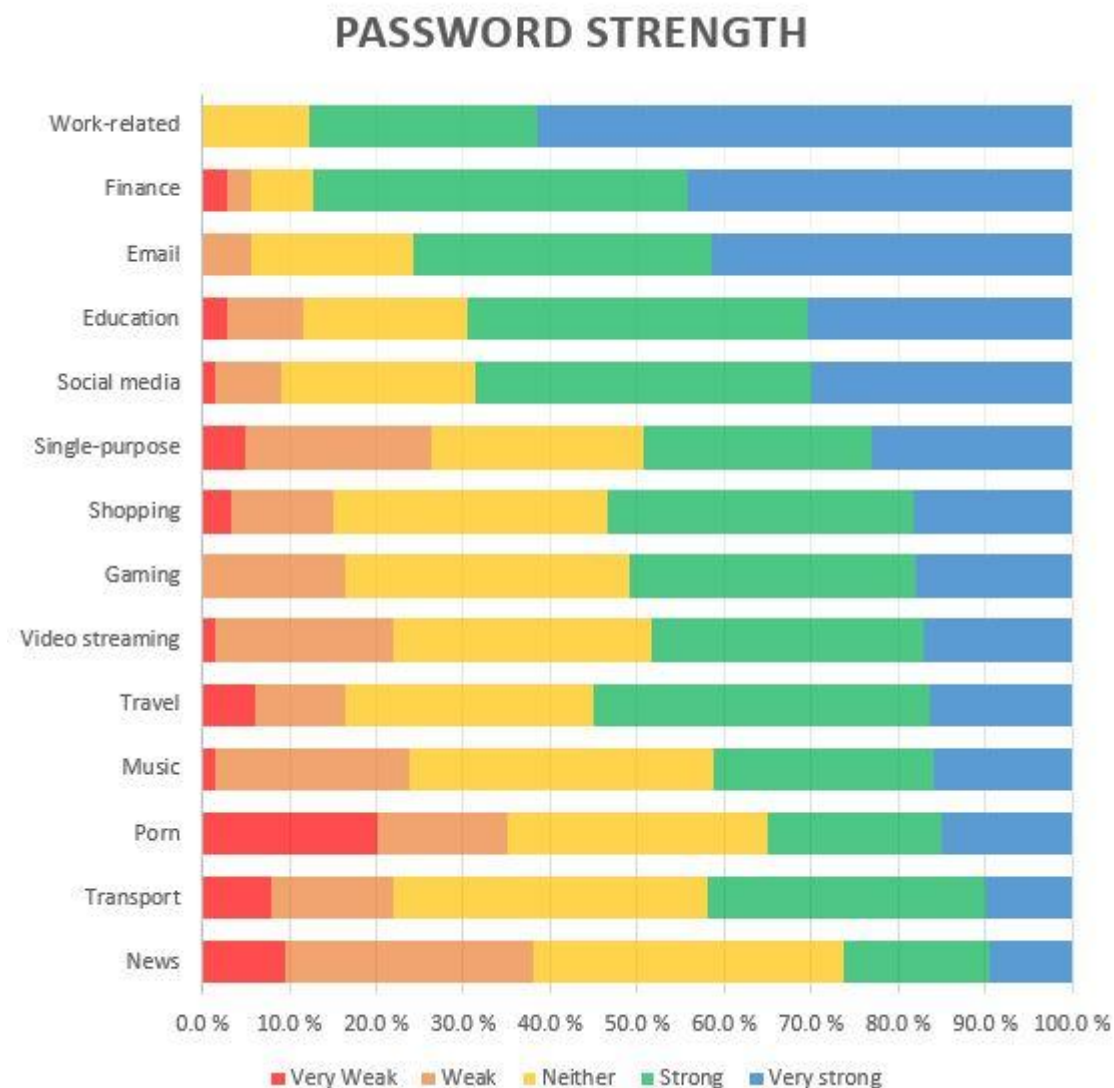


Figure 3 Password strength in online services

Work-related (61.5% very strong) services are ranked the highest with a good margin compared to other services, as the one which the participants reported to have the strongest password overall, followed by financial services (44.3%) and email (41.4%). Education (30.4%) and social media (29.9%) gave almost the same results. Single purpose (23.0%), shopping (18.3%), gaming (18.0%), video streaming (17.2%), travel (16.3%) and music (15.9%) also yielded very similar results. Porn (15.0%) also outperforms transportation (10.0%) and news (9.5%). Those who did have an account differ greatly in their responses. At first glance it appears to be the services which could be described as “important” compared to porn, transport, and news have overall stronger passwords. Porn had the biggest spread in strength rating. Only a few admitted to having very weak password on their services, only work-related, email and gaming was exempt from this result.

The services which had an overall poor scoring in password strength resulted in news services (38.5% weak and very weak), porn (35.0%), and single purpose (26.2%).

4.3.2 Likelihood of compromise

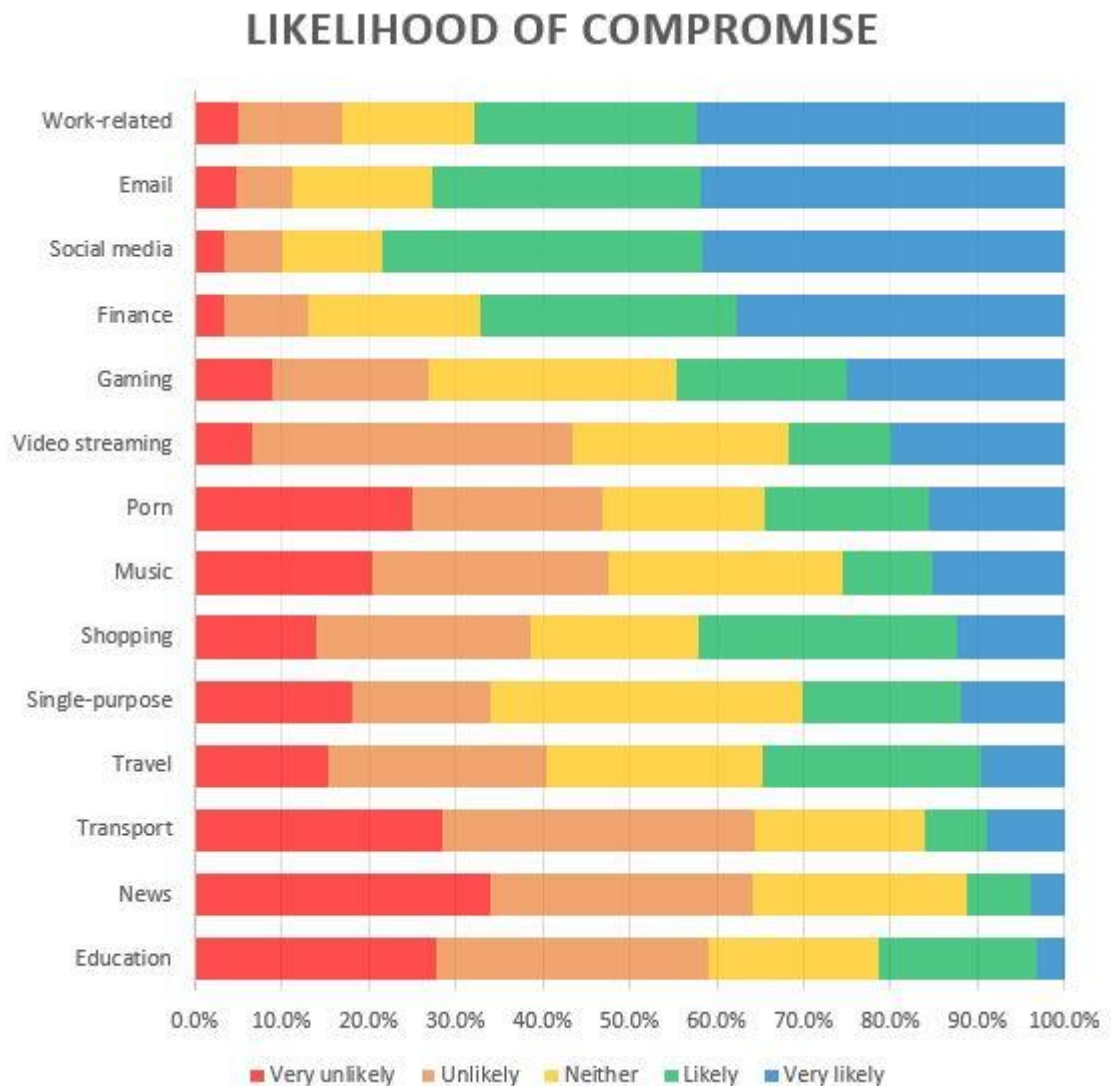


Figure 4 Perceived likelihood of compromise

Work related (42.4% very likely), email (41.9%) and social media (41.7%) are the categories which most users believe are the most likely for someone attempting to compromise. Close by was finance (37.7%) as they might have highest value for an attacker. Gaming (25%), video-steaming (20%), porn (15.6%) music (15.3%) and shopping (12.3%), single purpose (12%) and travel (9.6%) show significantly less likelihood in comparison and gives a more evenly distribution. Transport

(8.9%), new (3.8%) and education (3.3%) were perceived to be less likely to be attempted compromised.

4.3.3 Importance of availability

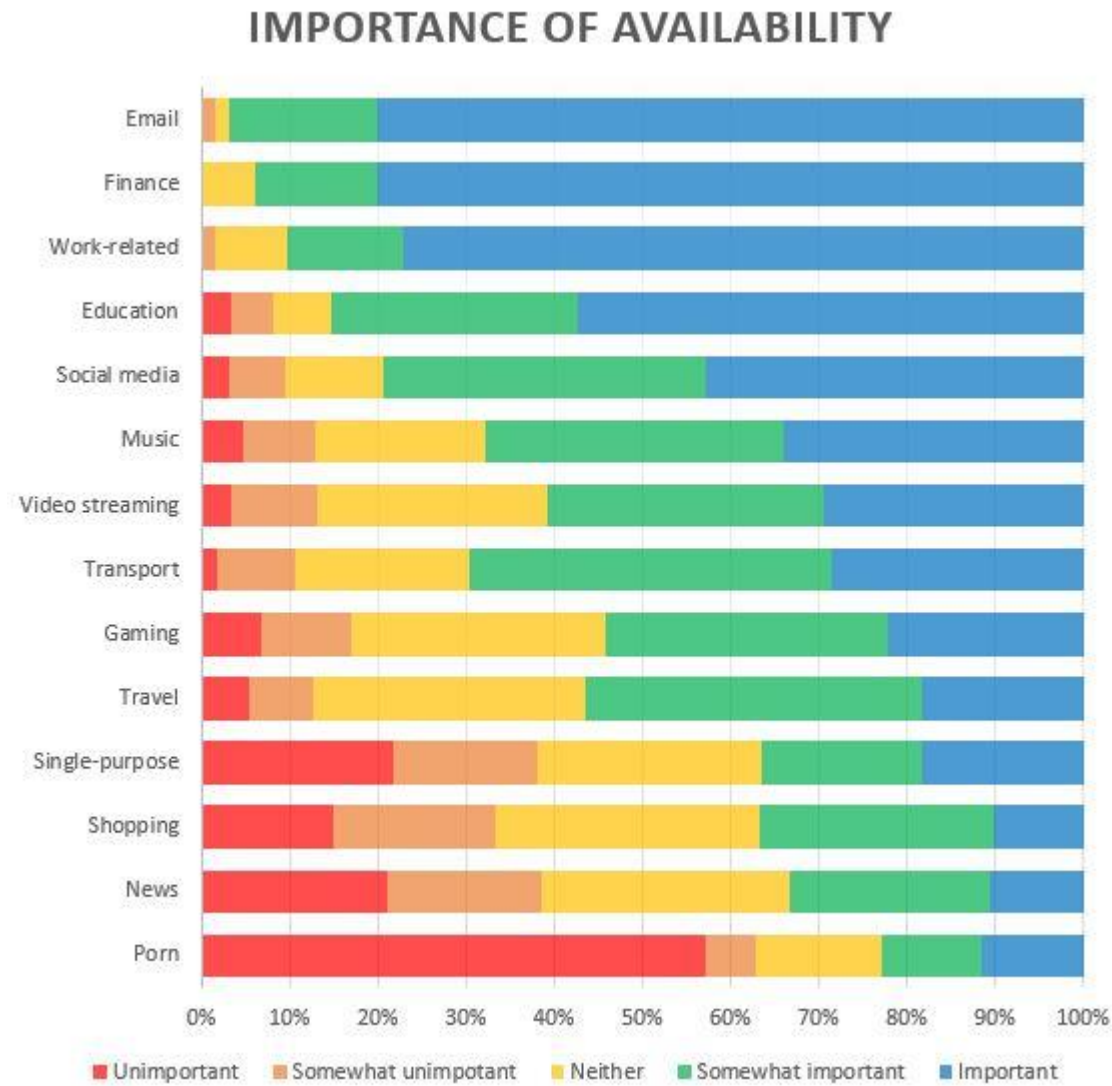


Figure 5 Importance of availability

Email (80.0%), finance (80.0%) work related (77.0%) services were perceived the most important regarding availability, perhaps because of how integral they have become to everyday life, followed by education (57.4%) which comes as no surprise as the participants were students during the survey's duration. Social media (42.9%), music (33.9%), video streaming (29.5%), transport (28.6%), gaming (22%) and travel (18.2%) are even less important. Single purpose (18.2%)

was the most undecided as it had the highest standard deviation amongst the service, similar were porn (11.4%), news (10.5%) and shopping (10%).

4.3.4 Consequences

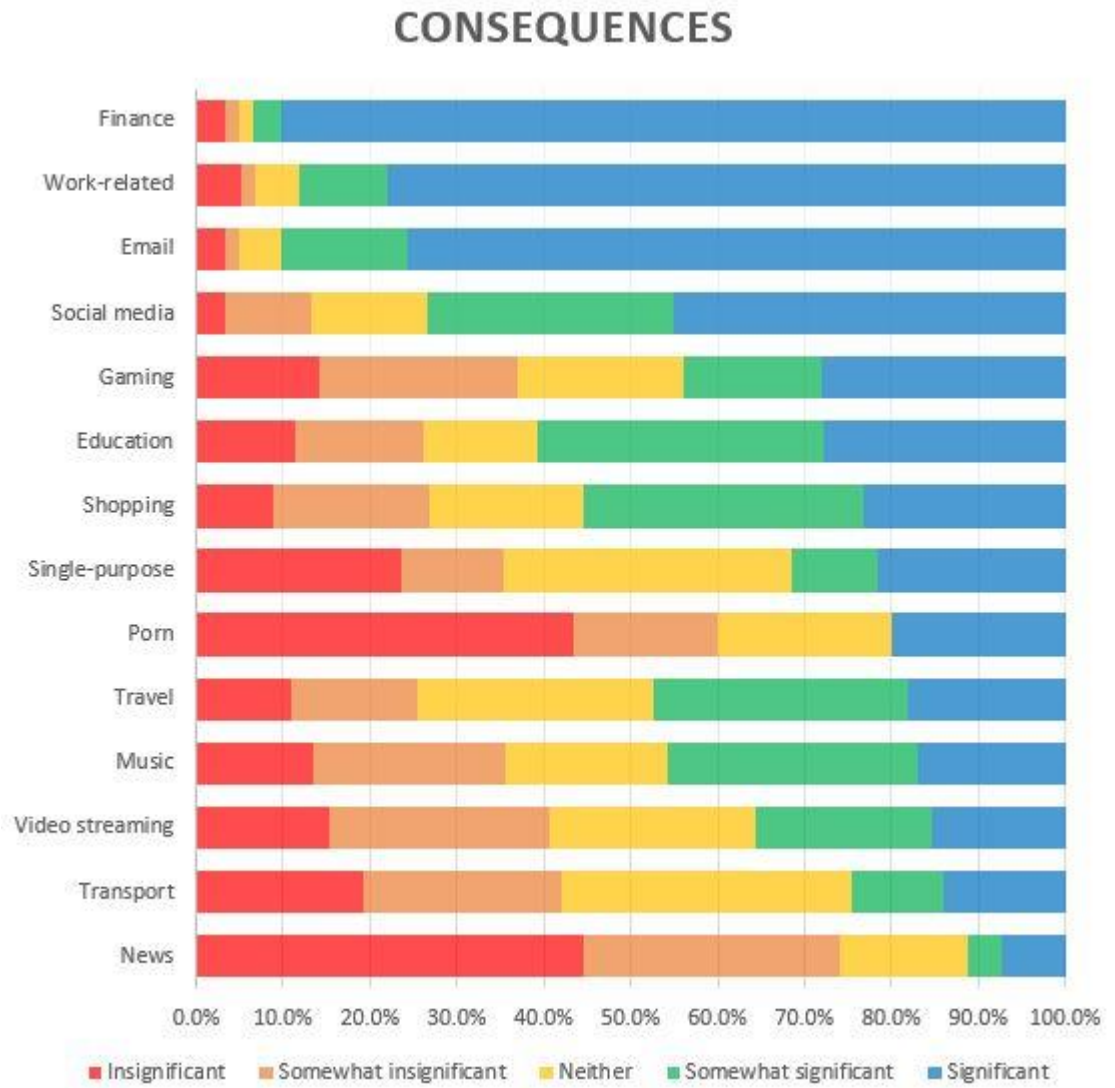


Figure 6 Perceived consequences

Finance services (90.2%) is the one with highest perceived consequences of a compromise, followed by work related (78%) and email (75.8%) which gives an almost unanimous response. Social media (45%) is not as significant as the top three but is still ahead of the curve. Gaming (28.1%), education (27.9%), shopping (23.2%), porn (20%), music (16.9%), travel (18.2%) and video streaming (15.3%) have equivalent results. Single purpose (21.6%) has the highest standard deviation.

News (7.4%) is the service which overall has the lowest consequences with the majority reporting it as either insignificant or very insignificant.

4.4 Comparing men and women

When comparing the sexes in the different services we excluded those who either did not give an answer or chose not to say (1.4%). The data is presented as the average answer given by each sex where each answer have an associated score; stronger answer give higher scores. The graphs are adjusted to give a more digestible impression. It is sorted in descending order from the men's highest to lowest with the x-axis range narrowed to give a better nuance. The graphs' range is scaled to emphasize differences.

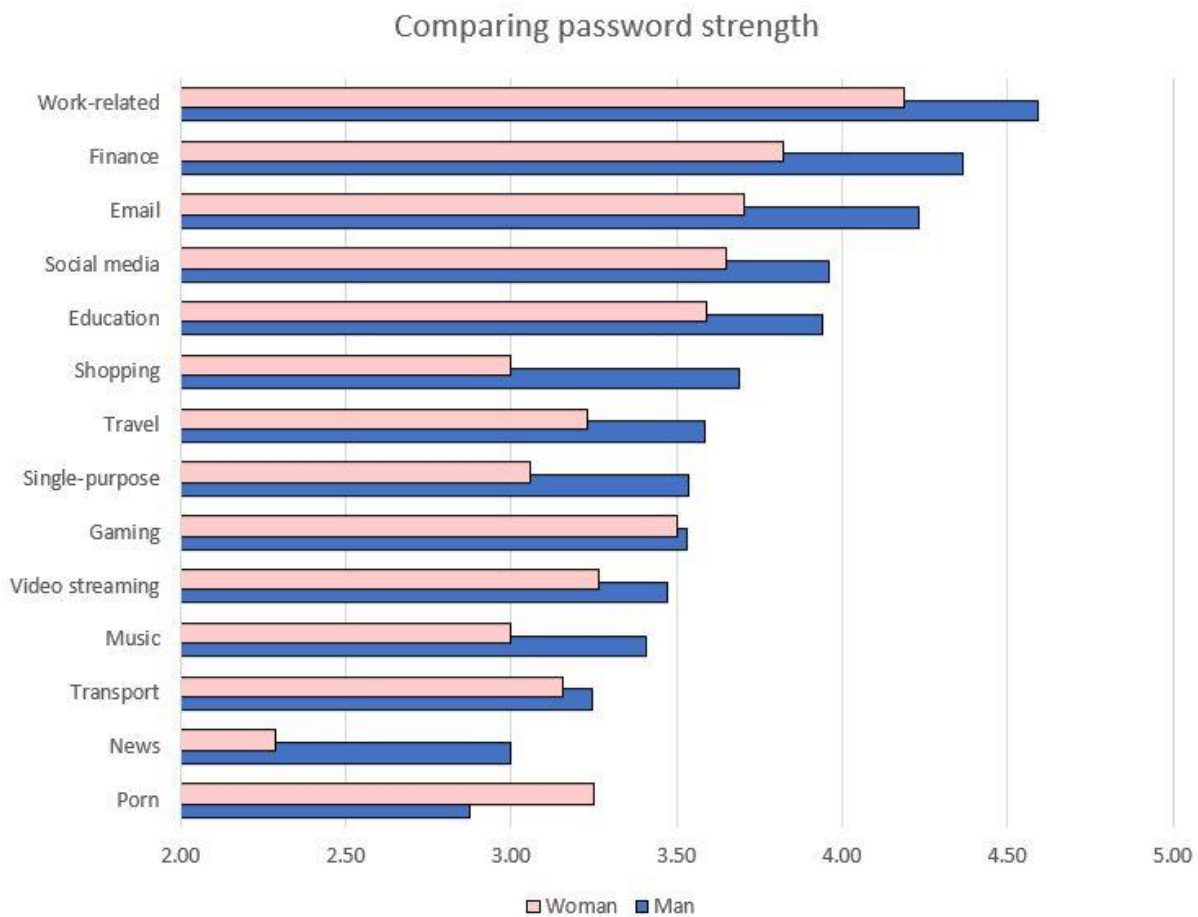


Figure 7 Comparing password strength distribution between men and woman

Men have a higher average score on password strength although the differences are very small. Men had an average score of 3.67 and women 3.34. Work related accounts were the highest scoring for both men and women. The only category

men and women completely agreed on was gaming. Women only scoring highest compared to men in porn services where they reported it as having slightly stronger password. Porn was also the service which men reported as having the weakest password. Women reported that news services had the overall lowest score of password strength.

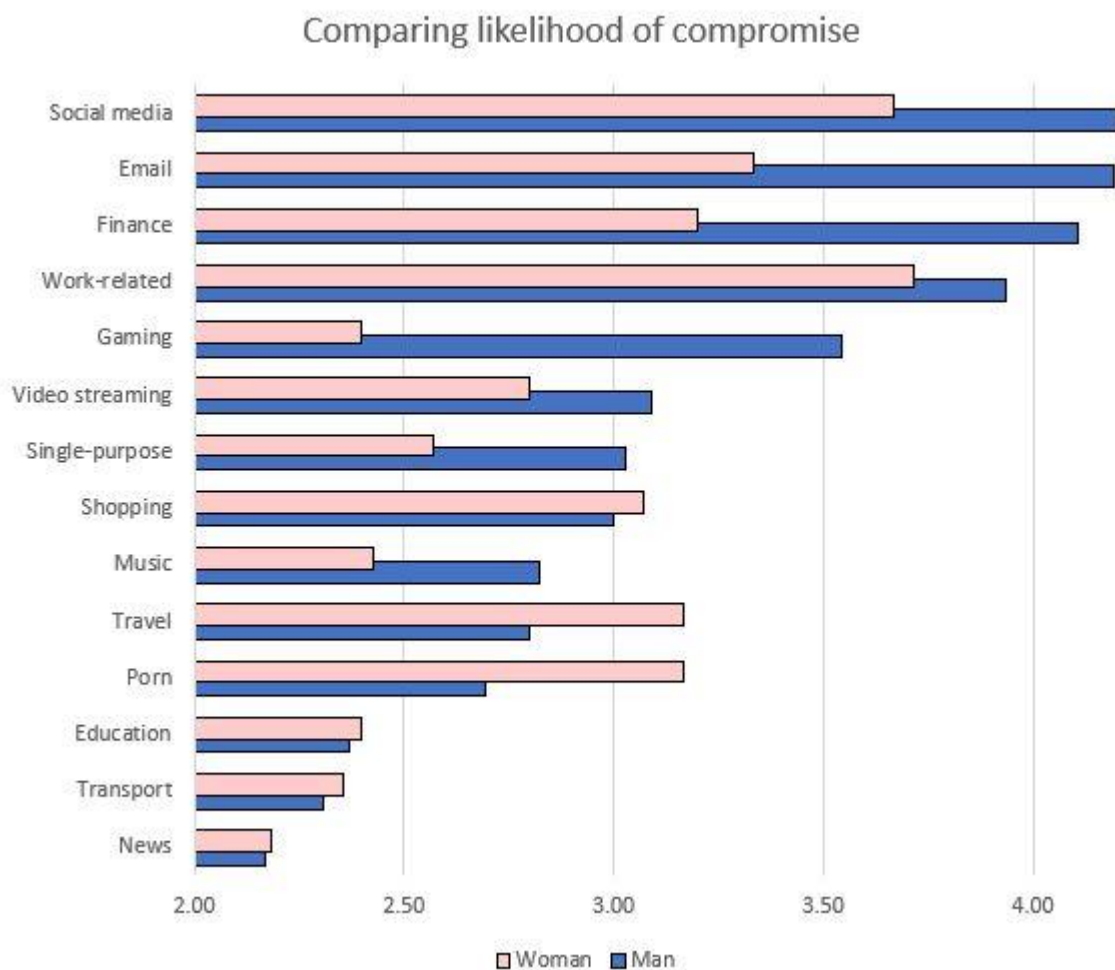


Figure 8 Comparing perceived likelihood of compromised between men and women

Men have a higher average score when perceiving likelihood of compromise on the different services (men: 3.17 and women: 2.89), presenting men as comparatively more careful than women. Social media, email, finance, and work related are all high scoring for both men and women. Men scored significantly higher on gaming services than women. A consensus was reached in shopping, education, transport, and news as both judged them as evenly likely to be attempted compromised. Women scored higher on travel and porn. We are unsure as to why travel have a higher score amongst woman, but porn can probably be explained by

the social climate with woman and sexuality as it still might be tabooer for woman than men.

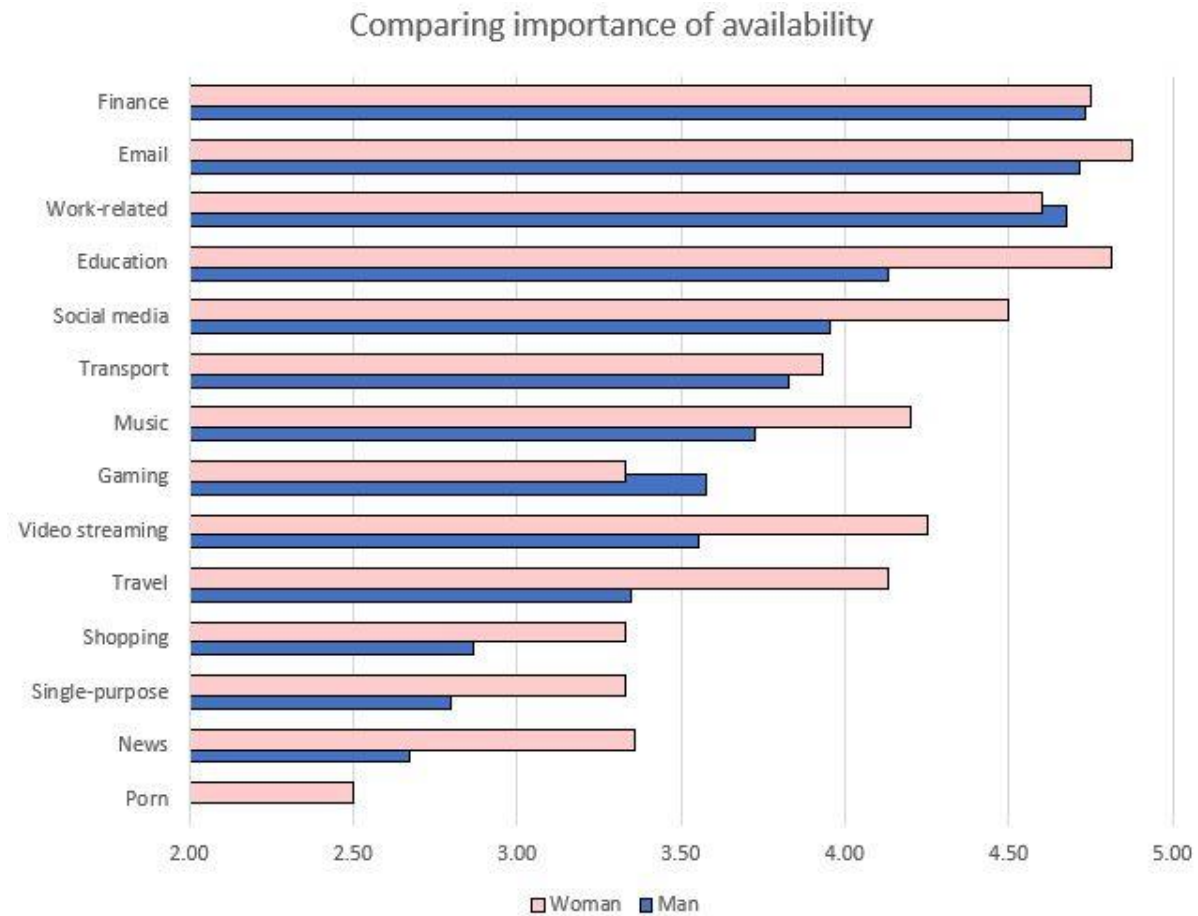


Figure 9 Comparing importance of availability between men and woman

Women scored a higher average on the importance of availability (men: 3.61 and women: 3.99). Both sexes agrees that finance, email, and work related are the services most important to have available, it comes as no surprise that these services are equally important as they neither are associated with fun or pleasure, but rather daily necessity. Men scores relatively low in porn. Women surpasses men in every category except gaming, albeit barely, and especially in porn, travel, video streaming, news, and education. This makes the man seem more carefree when it comes to non-critical online services.

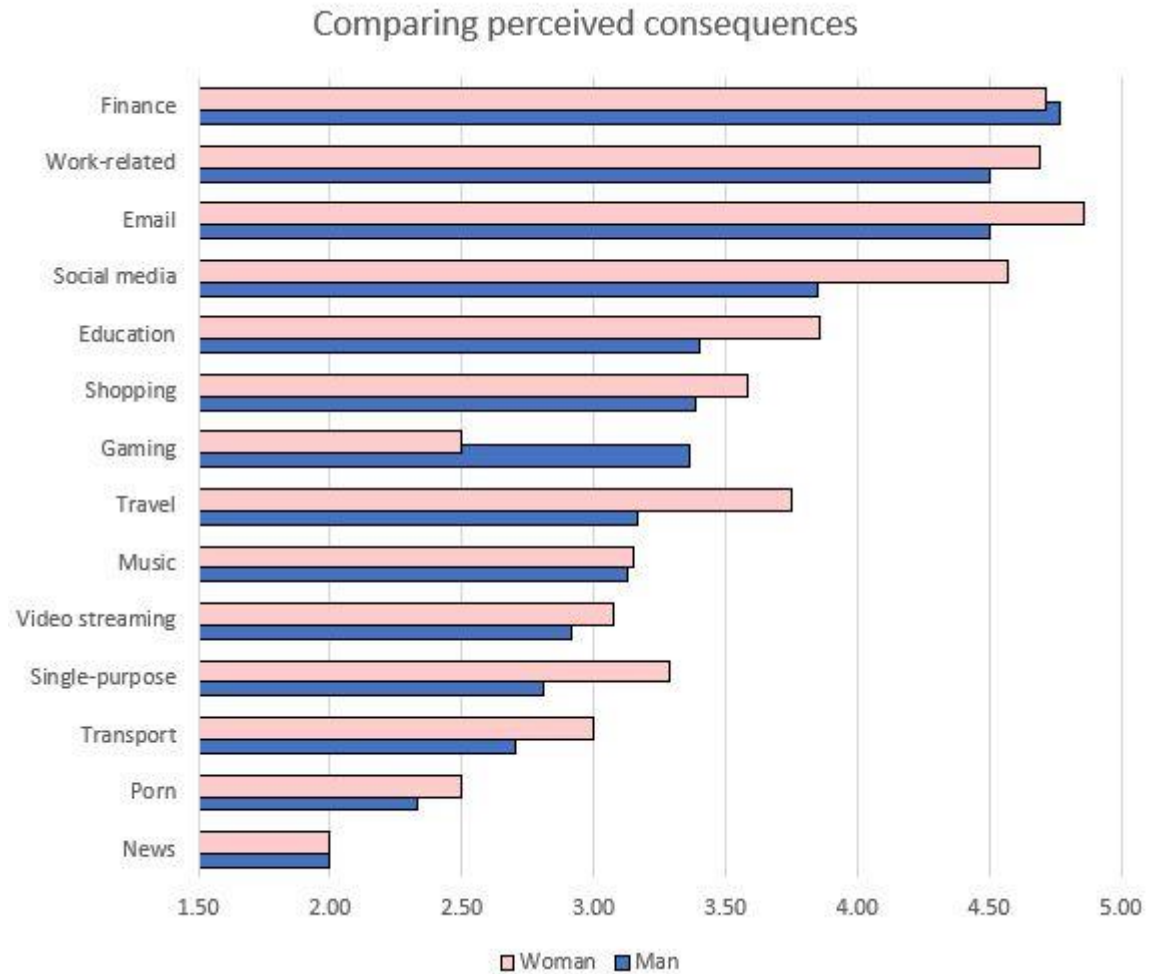


Figure 10 Comparing the perceived consequences of a compromise between men and woman

Both men and women score similarly (men: 3.34 and women: 3.54) and agreeing on finance, work related, and email would have the most significant consequences in a data breach. Women scoring slightly higher overall except for the disparity in gaming services were men reported notably higher consequences. Women outperforms men when perceiving consequences in social media and travel.

4.5 Summarize results

To summarize the survey findings, we would first like to point out that the porn category only had 21 respondents (30.0%). This means that the results presented to this category carries the risk of being unrepresentative. What it does show is that

most participants do not have a user account on porn sites or was just unwilling to answer the question.

When it comes to how password strength was distributed amongst the different services, our study clearly shows that work related accounts had the strongest passwords. Finance, email, education, and social media is also in a league of their own. Users also report low usage of weak or very weak passwords in every service except porn, news, and single purpose.

Importance of availability and consequence of compromise had the biggest difference between the services. Email, finance, and work-related services having a significantly higher score than other services. Porn and news had a low score for both sexes.

Likelihood of compromise had the most evenly distributed score amongst the services. In general, users found it unlikely or even unlikely that an attacker would try to compromise their account for most services, except for work-related, email, social media and finance being the outliers to this generalization for being.

To summarize the findings comparatively to the sexes it showed that men and women's perceptions of online services appears to be very similar in their responses, with only small overall differences ranging between 5-10%. Men are more consistent in their risk perception in gaming, whilst women are most consistent in shopping, travel, and porn. There are a few examples where there are significant differences in password strength and perceived risks. Porn being the one exception as it is perceived to have higher risk by woman, and woman use stronger passwords on this online service.

5 DATA ANALYSIS

In this chapter we'll investigate the results from the survey-findings and perform correlations-analysis between different factors to find any connections and patterns. The analysis will be elaborated as tables and graphs are presented. The data here is only presented as the result of the analysis, it will be further discussed in the next chapter.

5.1 Users

By doing a brief correlation analysis of the demographic values against password strength and risk variables, we will quickly verify if there are any significant variables we would have to take into consideration.

Table 11 Correlating age and risk variables

	Password strength	Likelihood of compromise	Importance of availability	Consequence
Age	-0.0157	0.0490	0.2094	0.1178

Out of the 70 respondents only 63 included their age in the survey, and only those who did are included in this table. When correlating the age and password strength it gave a negative negligible correlation coefficient of -0.0157, which means that in our sample age have nothing to do with password strength. It could be heavily influenced by the fact that 55% of those who reported their age is in the age group 24-26 years. Similar negligible results were given for the correlation coefficient of age and likelihood of compromise, as well as consequence. Importance of availability score slightly higher than the rest but is still negligible. To get a deeper understanding of user habits investigate the relation between the participants mean password strength and its corresponding standard deviation. This allowed us to observe the spread in password strength as it gets stronger between the online services for the individual users.

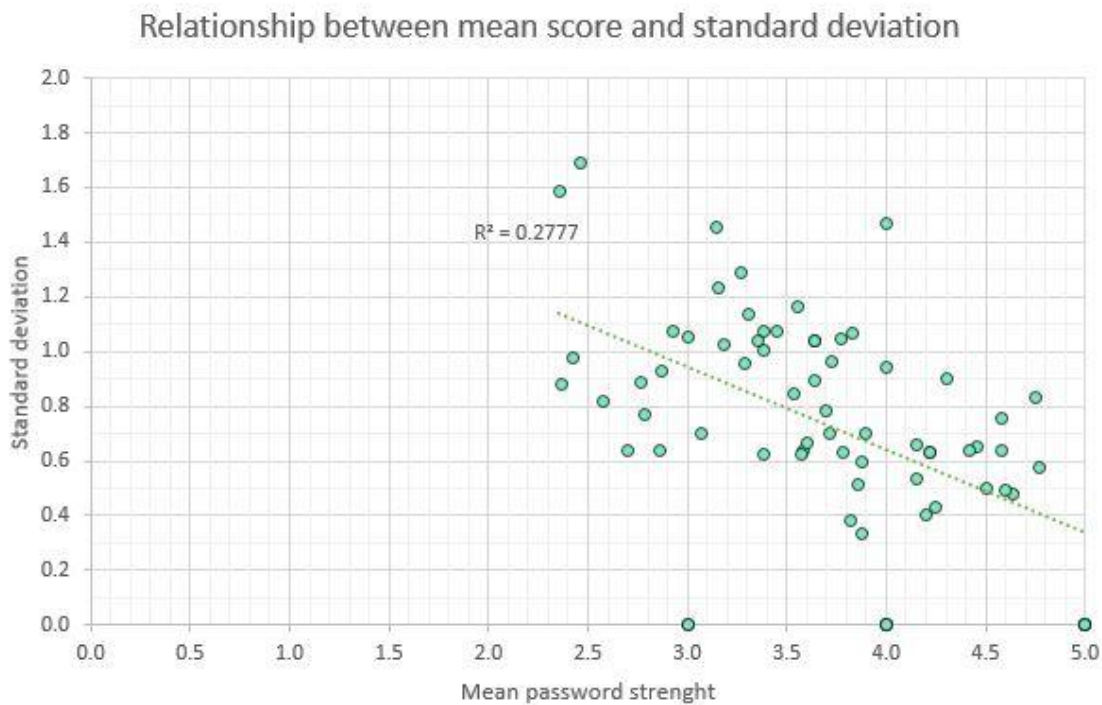


Figure 11 Relationship between mean password strength and standard deviation

When looking at how password strength compares to its corresponding standard deviation in this boxplot-graph. By looking at the trend line it is shown that as the higher mean password strength goes, the less deviations in password strength there is, which again could indicate that with strong passwords have tendency to have strong passwords on other online services as well. This could also be interpreted as those with a lower mean password strength score have a somewhat higher deviation, meaning they are more likely to vary more in their password strengths across multiple services.

The coefficient of determination, shown as R^2 represents the relationship between variable x and y in the form of a percentage. It is used in this case to predict password future password behavior by quantifying how close the variables are to each other compared to the trend line. This shows that there is only a 27.8% relation between password strength and standard deviation. This means that mean password strength only is accountable for 27.8% of the variation in standard deviation.

The outliers at the bottom are users who exclusively reported the same results across every online service, only fives, fours, and threes. The ones floating above are those to varied significantly higher in their answers compared to the other participants.

5.2 Risk factors

By testing some of the demographic variables we'll be able to get a result which could be used as a basis of comparison when comparing it with similar studies. We'll try to explain the different correlations they're graphically presented.

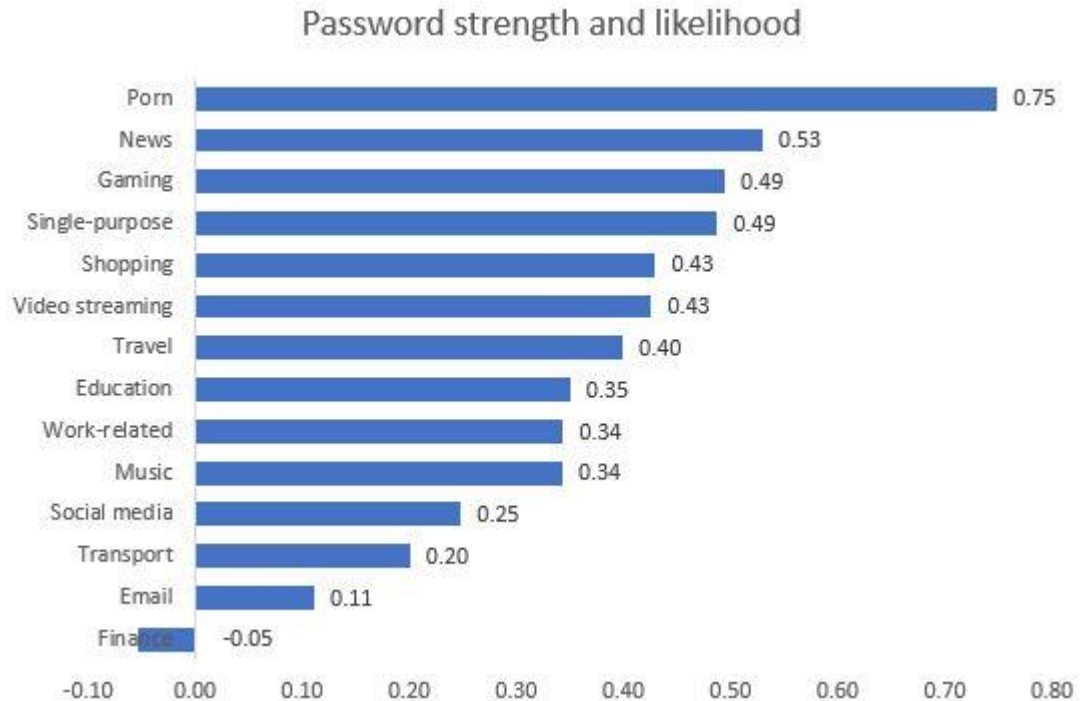


Figure 12 Correlation between password strength and likelihood of compromise

Looking at the correlation between password strength and the likelihood of someone attempting to compromise, we find that porn is the only service with a high correlation, news being the only one with moderate correlation, and most services only results in a low positive correlation with social media, transport, email and finance albeit negative giving a negligible correlation. This result can be interpreted as perceived likelihood of compromise has some on influencing password strength but finding correlation first and foremost with services which can be categorized as non-critical to everyday life. This is made clear when looking at the lower end of the graph: finance, email, and transportation – which can be considered as most crucial than porn, news and gaming.

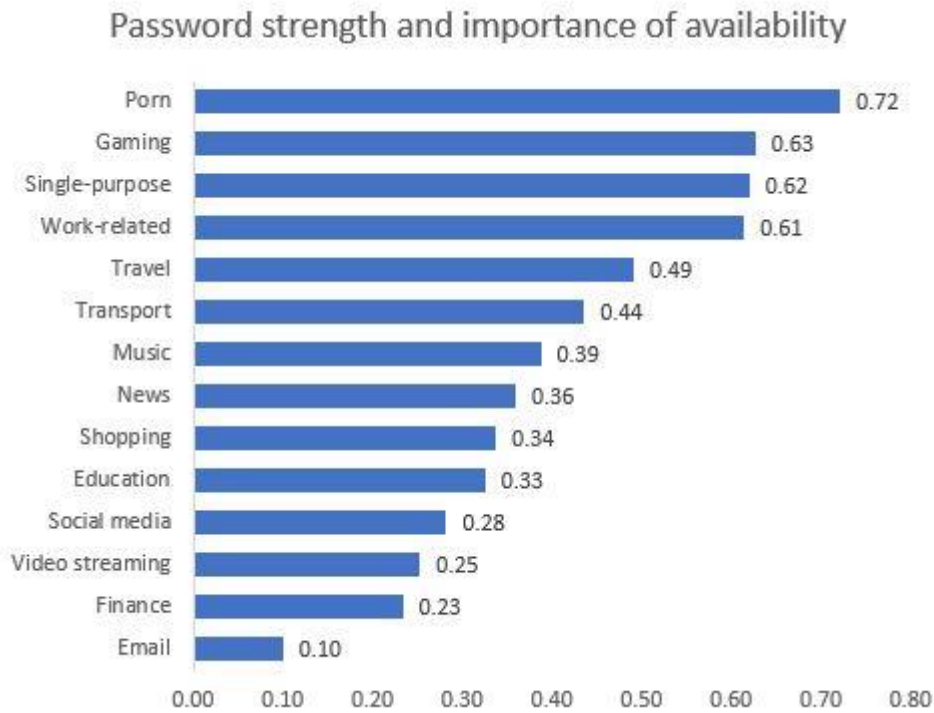


Figure 13 Correlation between password strength and importance of availability

When correlating services which availability is important with associated password strength, we find that there is a high correlation in porn services. Gaming, single purpose and work relate have a moderate correlation. Social media, video streaming, finance and email have a low correlation as the importance of availability affects overall password strength in a very small degree.

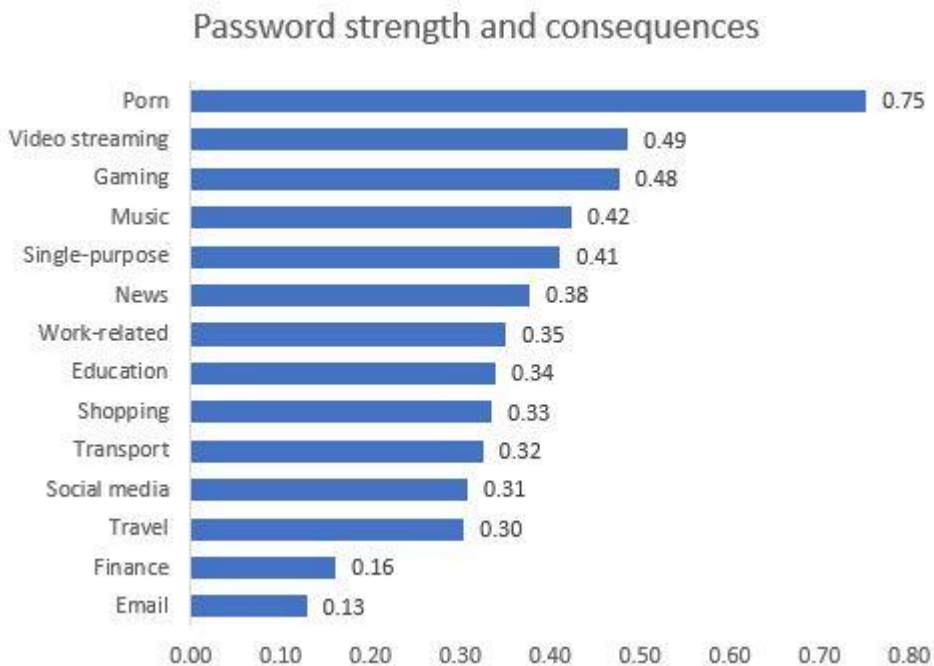


Figure 14 Correlation between password strength and consequences

Correlation between password strength gave a high correlation coefficient for porn as an outlier, and the rest of the services with a low correlation except for finance and email, which gave a negligible correlation. Trivial services have a much stronger correlation than essential services.

5.3 Online services

The respondents perceived the digital services differently, both in password strength as well as the risk factors, perception of likelihood of compromise, importance of availability and consequences. We'll explain each service how they're perceived and how the correlation can be interpreted.

5.3.1 Educational

Educational services were one the services which were applicable to every participant. This scored amongst the highest in password strength, 69.9% reported it as either strong or very strong. It was unlikely to be attempted compromised (21.3% said unlikely or very unlikely). It is one the most important services with 85.2% reporting it as either somewhat important or important. Over half the sample perceived consequences as 60.7% either somewhat significant or significant. These services have a low correlates with likelihood (0.35), importance (0.33) and consequences (0.34). It is understandable that students find educational services valuable as they have applied relative strong passwords. Understandable as the participants respondents educational accounts can be considered only valuable to them and thereby less valuable for a threat actor. The consequences are perceived as comparatively high as a compromise could cause complications in the form of denial of service.

5.3.2 Email

Email was reported as applicable to every participant, which comes as no surprise as it has become an ingrained service in everyday life. Although email services scores in the top of every category it yields very small correlations between password strength and risk factors. Its overall password strength is in the top three as 75.5% reported to have either strong or very strong, and very likely to be attempted compromised (67.8% reported it as either likely or very likely), 80.0% reported its

availability as important, and its consequences to be 75.8% significant. This gave a negligible correlation for coefficient with likelihood (0.11), importance of availability (0.10) and consequences (0.13). It appears as if this service is both important and crucial, but the mentioned risk-factors was of no concern when a password is created.

5.3.3 *Financial*

Financial services scored strongly in password strength as 87.1% reported it as having either strong or very strong credentials, and it is reportedly very likely to be attempted compromised, have the consequences compared to the other services and is second most important service to have available after emails. Even if the scores are overall very high, there is still a negligible correlation between strength and the risk variables. It has a -0.05 for likelihood, 0.23 for importance, and 0.16 for consequences. It is without a doubt a valued services as it warrants the strongest passwords and has the highest risks, but those variables can not explain the password strength.

5.3.4 *Gaming*

Gaming services scores relative average compared to the other online services with 50.8% reports to have either strong or very strong passwords, with 14.5% saying it's not applicable. It has a low correlation (0.49) with likelihood, a moderate correlation (0.63) with availability and a low correlation (0.49) with consequences. Which makes gaming the online service whose most influenced by risk variables. This is a good example of a service whose security reflects the risk perception of the end user, meaning for those who value it higher might give it a stronger password as its importance and consequences increases.

5.3.5 *News*

News services was rated as the service with the lowest scoring in password strength and highest responses for weak passwords. It is the service, which is less likely to be compromise, is the second to least important regarding availability, and has the lowest consequences if a data breach were to occur. It correlates moderately with likelihood (0.53), low correlation with importance (0.39) and

consequences (0.38). It became apparent that news related accounts are not important to students, this is reflected in the correlations. Taking into consideration that a news account does not necessarily limit your ability to read news articles

5.3.6 Porn

Porn is an outlier in the dataset because 49 respondents (71.0%) reported it as N/A not applicable, which means that the dataset is comprised of 21 respondents, which makes the data less reliable. The remaining data resulted in the service with the most normal distribution in both password strength and likelihood of compromise, albeit with the percentage of very weak passwords (20.0%). It is perceived almost as unlikely (46.9%) than likely (34.4%) to be compromised. It is the service which is deemed least important comparatively with 57.0% rating it's availability as unimportant, and its consequences in insignificant (43.3%). Its correlation coefficient is the overall highest amongst the services, most likely due to its low number of respondents, and therefore more volatile compared to other services. It resulted in high coefficient in likelihood (0.75), importance (0.72) and consequences (0.75). Due to the low sample size of "porn-users" it is hard to identify anything significant, but of those who did respond it was clear that perceived risk influences password strength. Riskier porn warrant stronger password.

5.3.7 Single purpose

Single purpose gave the most average responses as the risk variables were all normally distributed, it might be because it is the most ambiguous service in this thesis. It had a medium-strong password with 49.2% having either strong or very strong passwords. It is neither likely nor unlikely to be compromised, important nor unimportant, and its consequences is neither significant nor insignificant. Its correlation coefficient shows 0.49 with likelihood, 0.62 with importance, and 0.41 with consequences. To analyze this service is as ambiguous for the participants as it an enigma for us. This likely is because could be used as everything.

5.3.8 Shopping

Shopping is a service with monetary significance as transaction and credit information can be stored and used. The majority 53.3% reported it as having strong or very strong passwords. Perceived likelihood of compromising giving mixed results with 42.1% thinking it either likely or very likely, and 38.6% thinking its ether unlikely and very unlikely. This is also the case for the importance of availability as there is an almost equal perception as 36.7% thins its important (or somewhat important) and 33.3% thinks its unimportant (or somewhat unimportant). 55.4% also thinks its consequences are either somewhat significant or significant. Password strength gives a low correlation with likelihood (0.43), importance of availability (0.34), and consequences (0.33). These services could be influences by risk factors, but not to a degree that it could be generalized.

5.3.9 Social media

Social media is a popular service which represent a variety of different platforms, amongst the participants it was applicable to 97%. The majority (68.7%) reported its password strength as either strong or very strong, only 9.0% reported to having either weak or a very weak password. It is the service which is perceived to be most likely attempted compromised (78.3% either likely or very likely), it is also very important to have available (79.4% either somewhat important or important) with consequences almost as dire with 45.0% reporting is as significant. When correlating with likelihood, importance, and consequence, we achieved negligible 0.25, 0.28, and low 0.31 coefficient respectively. Which shows that perceived risk is almost absent in password creation.

5.3.10 Video streaming

Video streaming is service with a comparatively average password strength (48.4% either strong or very strong) with only 31.7% thinking it's likely or very likely some attempts to compromise these kinds of user accounts. Surprisingly its availability considered as important as 60.7% reported is as either important or very important, even though its consequences are on the lower end as 40.7% reported it as either somewhat insignificant or insignificant. It gives a negligible coefficient (0.25) with likelihood, low correlation with both likelihood of compromise (0.43) and consequences (0.49) Indicating that end users would prefer

the service it to be available, but if it for some reason was compromised it would not affect them too greatly.

5.3.11 Transportation

Transport is an everyday service, but not applicable for everyone as 27.9% reported it as not being. It has the highest number of respondents saying its password strength is neither (36%) strong nor weak. It is in the bottom tier when it comes to perceived likelihood of compromise with only 16.1% reported it as likely or very likely. In stark contrast is how important availability is with 69.6% saying its either somewhat important or important, even though consequences are low (24.6% either somewhat significant or significant). It has low correlation with likelihood (0.20), and a low correlation with consequences (0.32) and importance (0.44). This is one of those services which are crucial for those who rely on it and negligible for those do not.

5.3.12 Travel

Travel in contrast to transport is not necessarily an everyday service, but even still it was reportedly not applicable to 27.9%. Password strength was majorly strong as 55.1% said it had strong or very strong passwords. There were slightly more who thought it as more unlikely than likely that it would be attempted compromised (40.4% vs. 34.6% respectively). 56.4% think that availability is either somewhat important or important, which is as expected when comparing it to everyday transportation. Its consequences were reported as 47.3% either somewhat significant or significant, which makes it difficult to interpret as it is unclear if the participants think of consequences as a missed vacation or a data breach. It has a low correlation with likelihood (0.40), availability (0.49), and consequences (0.30). This could indicate that those who travel often also give the service slightly stronger passwords.

5.3.13 Work

Work related services have the overall strongest reported passwords and highest scoring risk values. It has an overwhelming strong password as 87.7% answered

either strong or very strong. It was reported to have the highest percentage of respondents perceiving it as very likely of compromise (42.4%), with 90.2% reporting its availability as either somewhat important or important. The same results were given in perceived consequences as 88.1% answered that would be somewhat significant or significant. Work related services has low correlation coefficient with likelihood of compromise (0.34) and consequence (0.35), and a moderate correlation with the importance of availability (0.63). There seem to be a strong consensus regarding risk perception and password strength, but its correlation with availability could indicate it as those who frequent this service more often and recognizes its importance might apply a stronger password to keep it so.

6 DISCUSSION

In this section we will discuss how the results and analysis of the survey answers the research questions specified in Chapter 1, and the limitations of this study including the sources of biases at the end. We start with the password strength distribution. Then move to risk perception and the correlation with password strength. Lastly, we address how the individual users behave. We also try to contextualize and compare our findings with past studies during each discussion section.

6.1 RQ1. Password strength

RQ1: To what degree is uniform password strength distributed amongst end-users?

When looking at chapter 4.3.1 and figure 3, we observe that password strength was not equally distributed amongst the online services as the participants in total reported having stronger and weaker passwords depending on the service. Based on our results, the services can be categorized into three different groups. Group one consisting of work-related, finance, email, education, and social media related services which clearly had the strongest reported passwords. Group two consisting of single-purpose, shopping, gaming, video streaming, music, travel, and transport related services. The third group consisting of music, porn, transport, and news, being the group with weakest reported passwords. As research in password behavior according to services is scarce, this finding cannot be contrasted with the literature. However, even though users generally reported to have strong passwords on most services, almost never used very weak passwords, and rarely weak passwords, this is not in line with reality (Taneski, Heričko, & Brumen, 2019; Juozapavicius, Brilingaite, Bukauskas, & Lugo, 2022; Grobler, et al., 2020). However, our findings could be supported by one study that states that users do not understand what constitutes a strong password (Ur, et al., 2016). Most users underestimate how many attempts a password should withstand, and struggle to differentiate between the strongest password when given two passwords to compare. As observed in chapter 4.2, figure 2, 64% of our sample group consist of students from either computer science, engineering, or social science, fields with

a seemingly higher likelihood of strong password usage. This is probably not enough to explain away our findings of users reported strong password usage. A study from 2021 found little significant difference between educational fields when investigating password strength (Salem, Moreb, & Rabayah, 2021).

When comparing men and women, our male participants reported using stronger passwords for every service except porn, as observed in chapter 4.4, figure 7 these findings are consistent with other studies concluding that men have slightly stronger password than women (Bonneau, 2012) (Mazurek, et al., 2013), this might be true overall, but not in every category as the women in our study reported stronger passwords in porn.

Thus, the degree in which uniform password strength is distributed among end users is to a very little degree as it varies according to the type of service which it is applied. There appeared to be only a minority who had reportedly, uniformly strong passwords.

6.2 RQ2: Perceived risk in digital services

RQ2: What is the correlation between password strength and perceived risk in digital services?

Figure 3, 4, and 6 in chapter 4.3, illustrate that close to 40% of our respondents perceive that most services are unlikely to be a target of attack, and that the consequences of a data breach would be either somewhat insignificant or insignificant. Compared to perceived password strength, where only 20% or less percent of users reported using weak passwords for most services, there is a disconnect between the perceived likelihood of a compromise, its consequences, and the strong password behavior. From our data we also observe, that the “neither” answer is more used in password strength, which could point to users not really considering their passwords as weak nor strong, or that they just don’t know. Furthermore, when comparing this with likelihood and consequence of compromise, where few users reported a likelihood or consequence of compromise as neither unlikely or likely, or insignificant or significant. In our study, perceived risk between some of the services has significant differences, especially when comparing the top three highest and lowest rated services as observed in figure 4, and 6 in chapter 4.3. This finding contradicts a similar study which found little difference between services (Merdenyan & Petrie, 2017). It is unclear why our findings differ. The difference could partially be explained by our number of added categories of services. If we compare the four categories used by Merdenyan and Petrine (SNS, email, eBanking and eComm) with our most similar categories

(social media, email, finance, and shopping), we do not see as much of a significant difference, but the difference between email and shopping in our dataset is still significant. Both when looking at likelihood of compromise and consequence of compromise, see figure 3 and 4 in chapter 4.3.

We found overall low correlation between both likelihood of compromise and password strength, and consequence of compromise and password strength as seen in figure 13 and 15 in chapter 5.2. Most services have similar “low” correlation rating between two given variables, but there seems to be a pattern for email and finance continuing to have no correlation between two variables. The only similar study we found however found no correlation between perceived consequence and behavior when looking at general password habits (Merdenyan & Petrie, 2017). However, it is only possible to compare the similar four categories. Email and finance evidenced no correlation, and social media and shopping had very low correlation when looking at password strength and consequence of compromise. Theirs and our findings are therefore very similar when only looking at these four categories.

The results indicates that the more trivial or inconsequential a service is, the more its password strength is influenced by risk variables, as show porn has strong correlations with risk factors in comparison to financial services which have negligible correlations.

6.3 RQ3. Behavior Patterns

RQ3: To what degree are end users consistent in their behavior when choosing password strength for different online services?

Our results prove that the differences in password strength on different online services are not necessarily dictated by fear of breach or compromise. This is shown due to the low correlation between password strength and different types of risk (likelihood, importance, and consequences). We believe that perception of the different services plays only a small part in creating strong passwords, and that password habits play a much bigger role in users’ actual behavior. As shown in chapter 5.1, figure 12, a higher mean password strength value decreased the standard deviation value. This could indicate that the users using strong passwords do not change their password strength much based on service, but rather uses the same general strength for all.

The findings of this study could also support previous studies on frequent password reuse (Rinn, Summers, Rhodes, Virothaisakun, & Chisnell, 2015; Salem, Moreb, & Rabayah, 2021). Moreover, the reason for users reporting similar

password strength perception across services could be explained by their password not changing much from account to account (Merdenyan & Petrie, 2022; Gratian, Bandi, Cukier, & Dykstra, 2018). One possible explanation could be that users maybe makes one unique password which they consider strong and reuse this password for several accounts with small to no changes for each reuse. Then they have another password which is reused on other accounts.

6.4 Limitations

We must acknowledge this study's limitations in terms of generalization, measurement and user behavior when answering the survey. We will present the main limitations of this thesis.

Some of the values in our study are ordinal, such as password strength and risk variables, because they are answered though a 5-point Likert scale. This means that we cannot accurately quantify the difference between each answer and thereby not get as a precise measurement as interval scale values.

Also, it is important to note that this study does not actually reflect the reality of users' password strength for each service, as we have not looked at their actual passwords for any of the services.

There are some general biases which can arise at different phases of the study while using the survey methodology (Shin, 2020; Bogner & Landrock, 2016).

Selection bias. For example, there are some general biases which can arise at different phases of the study while using the survey methodology (Shin, 2020; Bogner & Landrock, 2016). For example, *Selection bias*, which is our biggest limitation, and more specifically *sampling bias*, as we got less then desired respondents in the survey. Therefore, it could be perceived as our results are presumedly biased given that our respondents had similar education background, resulting in the other fields of education not being represented. This means that the result may not completely represent the intended population of UiA.

Omitted variables bias could also be listed as one of our limitations, because we could use more variables in the analysis such as more specific types of risks and knowledge. We chose to limit our scope to only three risk variables (likelihood, importance, and consequence) due to the time it would take to finish the survey. Our survey design was focused on the time required to complete the survey to be less than 10 minutes in order to attract more respondents

Recall bias. Could occur when respondents are ranking the different password strengths for each service. We already know users struggle to remember passwords, so users will most likely not remember their password while answering

the survey. They might answer based on mixed feelings of what they think the strength should be, and then control for what they think in reality their password behavior is.

Confirmation bias. Especially in our conclusions and interpretations of the data, and how we choose to present the data in tables and graphs.

Question-order effects. We tried to avoid this by asking users about their password's strength first, then about the perceived risks. We found it more likely that thinking about risks first would affect password strength, and less likely that password strength would affect perceived risk responses.

Response order. As there are many services we asked for each variable, the likelihood of respondents to simply the response process increases. This could result in respondents not evaluating other alternatives when answering. Could also be one of the reasons why users did not differ much in their responses.

Despite the limitations, our study contributed to exploring password distributions and how risk is perceived on different services.

7 CONCLUSION

7.1 Final remarks

The purpose of this thesis was to find any indications of a coherent behavior in how password strength is distributed amongst online services. Our intent was to identify behavior that could lead security risks if it was discovered to be significant neglect when comparing online services. Out of our categories of services, those considered to be significant differences in the overall password strength by the survey participants were financial, email, social media, work related, and education slightly outperformed other services comparatively. When investigating correlations between risk factors (likelihood of compromise, importance of availability, and consequences) and password strength it only produced low to moderate correlation. Even though participants reported having varying risk perception of online services it appears to only have a slight influence on password strength.

It is our understanding that end users not necessarily take the perceived risk into consideration when creating a password even if password strength varies for different online services, there could be other factors not described in this thesis which could be the starting point for future studies.

7.2 Future work

We hope that this thesis encourages future studies related to password strength to push the field of information security and social science further.

It is our belief that there are differences in digital services and thereby overall patterns in security, and due to its myriad of factors it is hard to pinpoint the most crucial factors in password strength. It is also possible to expand the research in this thesis in a more qualitative approach.

Similar studies could also consider different demographic groups to either verify our research, expand our categories into concrete services, and thereby presenting the nuances each service represent in password strength and behavior.

8 REFERENCES

- 451 Research. (n.d.). *451 Research: Love 'em or Hate 'em, Passwords Are Here to Stay*. Retrieved May 31, 2022, from ENZOIC: <https://www.enzoic.com/451-passwords-are-here-to-stay/>
- ADSelfServicePlus. (n.d.). *why passwords are here to stay, and how IT teams can stay in control of password management*. Retrieved May 31, 2022, from ManageEngine: https://www.manageengine.com/products/self-service-password/why-passwords-are-here-to-stay.html?source=pwd_ebk_blog
- Awad, M. A.-Q. (2016). Password security: Password behavior analysis at a small university. *IEEE*.
- Bhandari, P. (2022). *Qualitative research*. Retrieved from Scribbr: <https://www.scribbr.com/methodology/qualitative-research/>
- Bhandari, P. (2022). *Scribbr*. Retrieved from Questionnaire: <https://www.scribbr.com/methodology/questionnaire/>
- Bhandari, P. a. (2021). *What is quantitative research?* Retrieved from Scribbr: <https://www.scribbr.com/methodology/quantitative-research/>
- Bogner, K., & Landrock, U. (2016). *Response Biases in Standardised Surveys (Version 2.0)*. Leibniz: Gesis.
- Bonneau, J. (2012). The science of guessing: analyzing an anonymized corpus of 70 million passwords. *IEEE Symposium on Security and Privacy 2012*, 538-552.
- Cain, A. A., Edwards, M. E., & Still, J. D. (2018). An exploratory study of cyber hygiene behaviors and knowledge. *Journal of Information Security and Applications*, 36-45.
- Calculator*. (2022). Retrieved from Sample size calculator: <https://www.calculator.net/sample-size-calculator.html>
- Creese, S., Hodges, D., Whitty, M. T., & Jamison-Powell, S. (2013). Relationships between Password Choices, Perceptions of Risk and Security Expertise. *International Conference on Human Aspects of Information Security, Privacy, and Trust*. Duncan Hodges.
- Creswell, J. W. (2003). *Research Design: Qualitative, Quantitative, and Mixed Methods Approaches* (2nd ed.). SAGE Publications.
- Difi*. (n.d.). Retrieved from ID-porten: <https://eid.difi.no/nb/id-porten>
- Fredericks, D. T., Futch, L. A., & Thomson, K. (2016). Comparing Student Password Knowledge and Behaviour: A Case Study. *Proceedings of the Tenth International Symposium on Human Aspects of Information Security & Assurance* (pp. 167-178). HAISA.
- Google. (2019). *Online Security Survey*. Google.
- Gratian, M., Bandi, S., Cukier, M., & Dykstra, J. (2018). Correlating human traits and cyber security behaviour intentions. *Computers & Security* 73 (2018), 345-358.

- Grobler, M., Chamikara, M. A., Abbott, J., Jeong, J. J., Nepal, S., & Paris, C. (2020). The importance of social identity on password formulations. *Personal and Ubiquitous Computing (2021)*, 813-827.
- Israel, G. D. (1992). Determining Sample Size. *University of Florida, IFAS Extension*.
- Joshi, A. K. (2015). Likert Scale: Explored and Explained. *British Journal of Applied Science & Technology*, 396-403.
- Juozapavicius, A., Brilingaite, A., Bukauskas, L., & Lugo, R. G. (2022). Age and Gender Impact on Password Hygiene. *Applied Sciences*.
- Kennison, S. M., & Chan-Tin, E. (2020). Taking Risks With Cybersecurity: Using Knowledge and Personal Characteristics to Predict Self-Reported Cybersecurity Behaviors. *frontiers in Psychology*.
- Lambert, V. A. (2012). Qualitative Descriptive Research: An Acceptable Design. *Pacific Rim International Journal of Nursing Research*, 255-256.
- LastPass. (2019). *The 3rd Annual Global Password Security Report*. LastPass.
- Lyastani, S. G., Schilling, M., Fahl, S., Bugiel, S., & Backes, M. (2018). Studying the Impact of Managers on Password. *USENIX Security*. Retrieved from https://arxiv.org/pdf/1712.08940.pdf%22%20%5Ct%20%22_blank
- Mamonov, S., & Benbunan-Fich, R. (2018). The impact of information security threat awareness on privacy-protective behaviors. *Computers in Human Behavior*, 32-44.
- Mazurek, M. L., Komanduri, S., Vidas, T., Bauer, L., Christine, N., Cranor, L., . . . Ur, B. (2013). *Measuring Password Guessability for an Entire University*. Carnegie Mellon University, University of New Mexico.
- McMillan, R. (2012, January 27). *The World's First Computer Password? It Was Useless Too*. Retrieved May 31, 2022, from Wired: <https://www.wired.com/2012/01/computer-password/>
- Merdenyan, B. P. (2019). Perceptions of Risk, Benefits and Likelihood of Undertaking Password Management Behaviours: Four Components. *Human-Computer Interaction - INTERACT 2019*, 549-563.
- Merdenyan, B., & Petrie, H. 2. (2022). Two studies of the perceptions of risk, benefits and likelihood of undertaking password management behaviours. *Behaviour & Information Technology*.
- Merdenyan, B., & Petrie, H. L. (2017). Perceptions of the risks of password related activities. *Proceedings of British HCI 2017*. Sunderland: BCS Learning and Development. doi:10.14236/ewic/HCI2017.54
- Mwagwabi, F., & Jiow, J. H. (2021). Compliance with security guidelines in teenagers: the conflicting role of peer influence and personal norms. *Australasian Journal of Information Systems*.
- Nilsen, K. (2020). *NSM Risiko 2020*. Sandvika: Nasjonal Sikkerhetsmyndighet.
- Ouytse, J. V. (2021). The prevalence and motivations for password sharing practices and intrusive behaviors among early adolescents' best friendships – A mixed-methods study. *Telematics and Informatics*.
- Parsons, K., Butavicius, M., McCormac, A., & Calic, D. (2016). Assessing information security attitudes: a comparison of two studies. *Information & Computer Security*, 228-240.
- PST. (2022). *Nasjonal trusselvurdering 2022*. Politiets sikkerhetstjeneste.

- Renaud, K., & Zimmerman, V. (2017). Enriched Nudges Lead to Stronger Password Replacements ... but Implement Mindfully.
- Rinn, C., Summers, K., Rhodes, E., Virothaisakun, J., & Chisnell, D. (2015). Password Creation Strategies Across High-and Low-Literacy Web Users. *ASIST*. St. Louis: ASIST.
- Robinson, K. (n.d.). *what is two factor authentication 2fa*. Retrieved May 31, 2022, from Twilio: <https://www.twilio.com/docs/glossary/what-is-two-factor-authentication-2fa>
- Rodríguez-Priego, N., Bavel, R. v., Vila, J., & Briggs, P. (2020). Framing Effects on Online Security Behavior. *frontiers in Psychology*, 11. doi:10.3389/fpsyg.2020.527886
- Salem, Y., Moreb, M., & Rabayah, K. S. (2021). Evaluation of Information Security Awareness among Palestinian Learners. *20 21 International Conference on Information Technology (ICIT)*. ICIT.
- Schaik, P. v., Renaud, K., Wilson, C., Jansen, J., & Onibokun, J. (2020). Risk as affect: The affect heuristic in cybersecurity. *Computers & Security* 90 (2020).
- Scopus. (n.d.). *About*. Retrieved May 31, 2022, from Scopus: <https://blog.scopus.com/about>
- Security.org Team. (2021). *America's Password Habits 2021*. Security.org.
- Seitz, T. H., Pfab, J., & Souque, S. (2017). Do Difference in Password Policies Prevent Password Reuse? *ACM CHI 2017* (pp. 2056-2063). Denver: ACM.
- Shin, T. (2020, February 18). *What is statistical bias and why is it so important in data science?* Retrieved June 2, 2022, from Medium: <https://towardsdatascience.com/what-is-statistical-bias-and-why-is-it-so-important-in-data-science-80e02bf7a88d>
- SpyCloud. (2021). *2021 Annual Credential Exposure Report*. SpyCloud.
- Stanton, J. M., Stam, K. R., Mastrangelo, P., & Jolton, J. (2005). Analysis of end user security behaviors. *Computers & Security*, 124-133.
- Steinbart, P. J., Keith, M. J., & Babb, J. (2016). Examining the Continuance of Secure Behavior: A Longitudinal Field Study of Mobile Device Authentication. *Information Systems Research*, 219–239.
- Szumski, O. (2018). Cybersecurity best practices among Polish students. *Procedia Computer Science* (pp. 1271–1280). Elsevier Ltd.
- Tam, L., Glassman, M., & Vandenwauver, M. (2010). The psychology of password management: a tradeoff between security and convenience. *Behaviour & Information Technology*, 233-244.
- Taneski, V., Heričko, M., & Brumen, B. (2016). Password security – no change in 35 years? *MIPRO* (pp. 1360-1365). Opatija, Croatia: IEEE.
- Taneski, V., Heričko, M., & Brumen, B. (2019). Systematic Overview of Password Security Problems. *Acta Polytechnica Hungarica* Vol. 16, No. 3, 2019, 143-165.
- Tarwireyi, P., Flowerday, S., & Bayaga, A. (2011). Information Security Competence Test with Regards to Password Management.
- UiA. (2022). Retrieved from Fakta om universitetet og dets historie: <https://www.uia.no/om-uia/fakta-om-universitetet-og-dets-historie>
- UiA. (2022, January). Retrieved from Likestilling, inkludering og mangfold på UiA: <https://www.uia.no/om-uia/likestilling-integrering-og-mangfold-paa-uia>

- Ur, B., Bees, J., Segreti, S. M., Bauer, L., Christin, N., & Cranor, L. F. (2016). Do Users' Perceptions of Password Security Match Reality? *ACM CHI 2016*. San Jose: ACM.
- Vekua, U. (2021, July 5). *types of authentication methods*. Retrieved May 31, 2022, from Veriff: <https://www.veriff.com/blog/types-of-authentication-methods>
- Verizon. (2022). *2022 Data Breach Investigations Report*. Verizon.
- Webster, J., & Watson, R. T. (2002). Analyzing the Past to Prepare for the Future: Writing a Literature Review. *MIS Quarterly Vol. 26 (2002)*, 13-23.

APPENDIX

Appendix A Survey



We appreciate you taking your time to fill this survey. It takes approximately 9 minutes to complete.

Disclaimer

Answers will be given anonymously, we will not collect or store personal data, and the results will be aggregated for analysis and deleted after the thesis' presentation 15.06.2022

This survey will ask you about demographic variables and attitudes towards different digital services and regards to passwords, risks, importance and consequences.

The data gathered in this survey will be used in the a cybersecurity masters thesis at University of Agder.

Sex

- Male
- Female
- Prefer not to answer

Age

Field of education

- Health
- Social care
- Sports
- History
- Philosophy
- Religion
- Engineering
- Computer science
- Fine arts
- Education
- Natural science
- Social science
- Economy and administration
- Law
- Linguistics
- Literature
- Communication and media

Years completed of higher education



0

10

Have you ever been hacked, or subjected to a data breach

- Yes
- No
- Dont know

A password manager is a software application designed to store and manage online credentials
Are you using a password manager?

- Yes
- No
- Previously

Consequences refers to the potential outcome if someone unauthorized accessed, disclosed, edited or deleted your data

Indicate how severe the consequences of a compromise be for you?

	Insignificant	Somewhat insignificant	Neither	Somewhat significant	Significant	N/A
Education (Canvas, Fronter, ItsLearning)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Email (gmail, hotmail, outlook)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Finance (online bank, investments)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Gaming (Steam, EpicGames)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Music (Spotify, Apple music)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
News (VG, Dagbladet, Dagens næringsliv)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Porn (Adult entertainments)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Shopping (Finn, Zalando, Komplet)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Single purpose (services which requires a user account)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Social media (facebook, snapchat, tiktok)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Transportation (public transport; bus, train)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Travel (Hotel, Fly, Booking)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Video streaming (Netflix, Youtube, HBO)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Work-related (requires a company login)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Thank you for completing this survey.

If you have any concerns or questions, please contact:

Håkon haakob16@uia.no
 Ole-Martin olemartinm@student.uia.no